A LINE-OF-SIGHT SENSOR NETWORK FOR WIDE AREA VIDEO SURVEILLANCE: SIMULATION AND EVALUATION

THESIS

Jamie R. Morrison, First Lieutenant, USAF

AFIT/GCE/ENG/05-05

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the U.S. Government.

AFIT/GCE/ENG/05-05

A LINE-OF-SIGHT SENSOR NETWORK FOR WIDE AREA VIDEO SURVEILLANCE: SIMULATION AND EVALUATION

THESIS

Presented to the Faculty

Department of Electrical and Computer Engineering

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science in Computer Engineering

Jamie R. Morrison, BS

First Lieutenant, USAF

March 2005

AFIT/GCE/ENG/05-05

A LINE-OF-SIGHT SENSOR NETWORK FOR WIDE AREA VIDEO SURVEILLANCE: SIMULATION AND EVALUATION

Jamie R. Morrison, BS

First Lieutenant, USAF

Approved:

/SIGNED/	16 MAR 05
Dr. Gunasekaran S. Seetharaman (Chairman)	Date
/SIGNED/	16 MAR 05
Dr. Michael L. Talbert, Lt Col, USAF (Member)	Date
/SIGNED/	16 MAR 05
Dr. Paul R. Havig (Member)	Date

Acknowledgments

I offer my sincere thanks and appreciation to my adviser, Dr. Guna Seetharaman. His visionary approach, thoughtful guidance, and understanding of the innovative process have proven invaluable to me in understanding not just how to pose solutions but to respect the breadth of possible solutions and bring the best of them to fruition. I also thank Lt Col Michael L. Talbert for his unwavering support. He stood as a constant reminder to enjoy doing the research; I don't think I've ever had so much fun working so hard. I thank Dr. Paul Havig for giving me a quiet space to work and sharing my interest in display technologies. I thank Dr. Rusty O. Baldwin for his advice, encouragement and insightful comments on the final draft of this document. Together they have all coached me pursue research in an exciting area. I am responsible for any unintended shortcomings that remain.

I thank my wife for her steadfast support in my studies and for taking such good care of our children over the many days that I could not be there with her. I truly believe that she has worked harder for this achievement than I have.

Most importantly, I give thanks to the glorious name of the Lord God. My discovery of Him has been the greatest revelation of my life. He has both saved me and sustained me such that any good thing I may have done is only by His grace.

Jamie R. Morrison

Table of Contents

Page

List of Figures

Page

List of Tables

Page

Abstract

Substantial performance improvement of a wide area video surveillance network can be obtained with addition of a Line-of-Sight sensor. The research described in this thesis shows that while the Line-of-Sight sensor cannot monitor areas with the ubiquity of video cameras alone, the combined network produces substantially fewer false alarms and superior location precision for numerous moving people than video.

Recent progress in fabrication of inexpensive, robust CMOS based video cameras have triggered a new approach to wide area surveillance of busy areas such as modeling an airport corridor as a distributed sensor network problem. Wireless communication between these cameras and other sensors make it more practical to deploy them in an arbitrary spatial configuration to unobtrusively monitor cooperative and non-cooperative people. The computation and communication to establish image registration between the cameras grows rapidly as the number cameras increases. Computation is required to detect people in each image; establish a correspondence between people in two or more images; compute exact 3-D position from each corresponding-pair; temporally track targets in space-and-time, and, assimilate resultant data until thresholds have been reached to either cause an alarm or abandon further monitoring of that person. Substantial improvement can be obtained with addition of a Line-of-Sight sensor as a location detection system to decoupling the detection, localization, and identification subtasks. That is, if the ‘where’ can be answered by a location detection system, the ‘what’ can be addressed by the video most effectively.

A LINE-OF-SIGHT SENSOR NETWORK FOR WIDE AREA VIDEO SURVEILLANCE: SIMULATION AND EVALUATION

I. Introduction

A. Motivation for Research

Video cameras have become increasingly powerful and less expensive, making them attractive for the monitoring and surveillance of large areas. Images from numerous cameras can detect, locate, identify, and track people for a location aware system. The task is to assimilate information in a concise manner and display it on a monitor for an observer, to facilitate an effective and timely response. It is desirable for the system to determine automatically whether to continue monitoring a person, or when to raise an alarm. In addition, it is desirable to be able to track nearly everyone until the person has been determined to be safe. However, the computation and communication requirements of such a system increase with the number of cameras and image resolution of the cameras. A difficulty in this type of system is to get the best possible identification image of people who are moving in the monitored area with a network of medium resolution video cameras.

Figure 1 illustrates the relationship between the number of cameras involved in determination of the location of people in an area. With few people in a small area, feature extraction can be used for location determination. Detection is the dominant source of complexity for this task. As the number of people and area for a location system increases, stereo image techniques become preferred. Location dominates the complexity in this domain. For numbers of people and areas become so large that they

cannot be effectively observed from a given vantage point, a video network becomes necessary. Difficulty of registering images and consistent labeling cause tracking to be the dominant source of complexity.

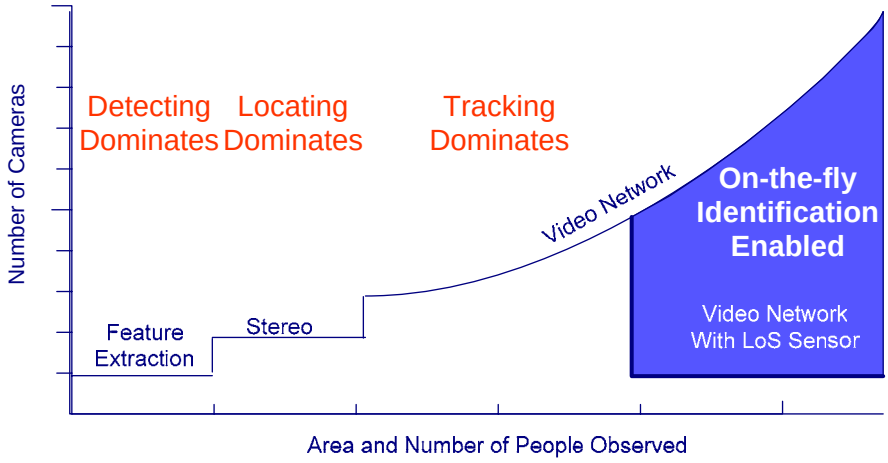


Figure 1. Number of Cameras to Identify and Track versus Number of People

A network of Line-of-Sight (LoS) sensors offers an efficient means of swiftly detecting and locating multiple moving people in a room or corridor. A Line-of-Sight sensor, by design, is a collection of light rays between a small number of light sources and light sensors opposite of each other in a fixed geometry, as explained later. Any object between the light sources and light sensors interrupts reception, enabling fast localization. Thus, a Line-of-Sight sensor is responsible for several light rays between sources and sensors. Because the network of Line-of-Sight sensors perform sensing of the environment in a distributed fashion, the location evidence the LoS sensor provides is

more concise than the location evidence derived by cameras. The video sensor network can further refine the already concise evidence with image data and use the results to perform identification and tracking. A further benefit of the LoS sensor may be obtained if the evidence from this sensor is used as a basis for registration communication between the cameras. Typically, video images have been used to answer questions such as: 'Is there something interesting?,' 'Where is it?,' 'What is it?' and 'What are its dynamics?' In principle, if LoS sensors can answer the first two questions, then, the video processing will be less demanding. The amount performance would increase due to this task decoupling will determined herein. From an implementation point of view, driven by enabling technologies, this approach seems promising. A performance evaluation may affect the design of future systems, which have been in demand since the September 11th events.

B. Experimental Hypothesis

The purpose of the multi sensor network is to locate all persons in the network's region of interest. The data from the sensor network determines the identification and track of each person detected. The LoS sensor is expected to provide a faster location detection of a person than three networked cameras, and it is expected the LoS sensor fused with any number of cameras will have fewer false alarms than three networked cameras.

C. Thesis Organization

Chapter one presents the overall organization of the thesis as well as a summary of the motivation for research and the research hypothesis.

Chapter two presents the background for this effort. A description of the role of sensor networks in location aware computing is given, as well as reasons for the popularity of the camera in such networks. A description of the LoS sensor is also given, including the algorithmic operation and complexity. Incorporated in this description is the motivation for feature level fusion of the sensor with the camera network along with a discussion of system deployment and scaling issues. Also included is a brief description of how the LoS sensor and camera systems contribute to each subtask of detecting, locating, identifying, and tracking people.

Chapter three presents the methodology for this effort. A simulation model of an airport corridor is created with eighteen sub-sections. There are three different types of subsections: a terminal with a desk, a terminal without a desk, and a portion of a corridor. Each subsection is equipped with three cameras and a LoS sensor configuration for estimating the location of people moving within the airport.

The measurements derived from the simulation include the hit rate, the false alarm rate, and the precision of the sensor network. The hit rate is the number of people in the simulation for whom the sensor network has correctly generated a corresponding location estimate. A false alarm is a location the sensor network has estimated a person is when no one occupies that area. The false alarm rate is the number of false alarms per

area per simulation. The precision corresponds to the number of pixels involved in the correct determination of a person's location.

The factors varied in the simulation include the number of people in the airport and the configuration of camera and LoS sensors.

Chapter four presents the results of the simulation. Because the purpose of the study is to evaluate the LoS sensor, the results are considered with respect to sensor network configurations that include this sensor and the configuration of three cameras networked without the LoS sensor. Analysis of variance is performed on the data. Confidence intervals of mean values and distributions of effects from factors and interactions between factors that produce a statistically significant effect on the variance are presented.

Chapter five contains the conclusions of the research and suggestions for future research. The ability of the LoS sensor to correctly and precisely locate a person is compared to that of a three-camera system. This chapter will show that the LoS sensor achieves better precision and produces fewer false alarms than the three-camera system, but has "blind spots" that prevent the LoS sensor from achieving the cameras detection ability. Finally, suggestions for further research are presented.

II. Literature Review

A. Introduction

The purpose of this chapter is to present the background for wide area surveillance using distributed sensor networks, and in particular, a hybrid multi-sensor network made of video cameras and specially designed LoS sensors. Important sub problems of location aware computing are introduced first to establish the context. Existing approaches are highlighted and the basic design of LoS sensor is presented. The suitability of the new sensor for detecting, locating, and tracking multiple people within the sensor-field is briefly explained. Our approach fuses video images obtained through a network of cameras, to the high-speed location sensing data through the line of sight sensors. The limitation of video-only networks is presented at a higher level to justify the need for augmentation. The scalability and performance of the combined approach is presented to form the basis for the proposed approach. A practical consideration in deployment of new sensors, among others, is wiring. Wiring is both expensive and intrusive of the existing infrastructure. Most distributed sensor networks, therefore, are likely to use wireless communication. Although wireless communication plays a vital role in the proposed framework, the scope of the analysis herein does not include these broader factors. However, the LoS sensor based fusion does minimize the information interchange between any pair of cameras sharing a common field of view, -- a desired feature in wireless sensor networks.

"The widespread deployment of sensing technologies will make location-aware applications part of everyday life. [13]" Context-aware computing, of which location

aware computing is a part, is the ability of an "information infrastructure" to respond to real-world context information such as a user's identity, current and recent physical location, weather, time, or user's dynamic state (sleeping, driving, walking, gestures) [22]. Location aware computing systems respond to past or present location information both spontaneously and on request. Mobile computing, a form of location aware computing uses wireless devices such as PDAs or laptop computers, to discover nearby resources such as printers or Wi-Fi "hotspots." Location aware computing also includes tracking and predicting location of objects including system users. Proposed uses of location aware computing include navigating unfamiliar spaces, locating office colleagues, tracking conference attendance, monitoring patients at medical facilities, and assisting the elderly or disabled [14][17]. Microsoft has announced that the soon to be released Longhorn operating system would include location-aware software components [13].

A fundamental problem for location-aware computing is determining physical location. The Global Positioning System (GPS) is a frequent choice for outdoor areas where five to ten meter accuracy is sufficient. However, GPS does not perform well indoors or in high-rise urban environments. The Active Badge system [20] uses infrared sensors arranged within a room to intercept infrared identification signals emitted periodically from badges worn by room occupants [29]. More accurate positioning was later achieved using radio systems such as Wi-Fi and Bluetooth technology. These systems provided location determination accuracy from tens of meters to within several meters. Ultrawideband radio, ultrasonic and computer vision technology systems now

provide location information from one-meter accuracy to within a few centimeters. Ultrawideband radio and ultrasonic systems use triangulation of signals received by several sensors to determine the location of badges or tags emitting the signals. The signals from such badges or tags typically contain identification information.

Computer vision offers attractive alternatives for monitoring and tracking people, since the subjects are not required to wear special badges or tags. However, simultaneous identification and concurrent tracking of numerous people in a dynamic environment is a significant challenge for a computer vision system [13]. This is especially problematic as the system must be efficient given the real-time operation of the system [18]. Identification performance can be improved by requiring users to wear tags with barcodes; this reintroduces the badge/tag requirement while also requiring system cameras have high enough resolution to read the barcode at expected distances [13].

Fusing image data from a camera in a computer vision system with other cameras or sensors at various fusion levels improves the systems ability to simultaneously identify and track numerous objects [12]. Table 1 compares the fusion of camera images at different levels. A system fusing camera image data with data from a sensor of different modality has been demonstrated [12]. The system locates people in a room through pixel-level fusion of images from a single camera using the differences in sound wave time-of-arrival data between a pair of microphones. This system is empirically demonstrated to locate two people speaking in a conference room, but performance metrics have not been determined [11].

Table 1. Comparison of Fusion Levels (from [12], [21])

Characteristics	Signal Level	Pixel Level	Feature Level	Symbol Level	Behavior Level
Type of sensory information	single or multi-dimensional signals	multiple images	features extracted from signals / images	symbol representing decision	motor action
Representation level of information	Low	Low to medium	Medium	High	Low
Model of Sensory Information	random variable corrupted by noise	stochastic process on image or pixels with multidimensional attributes	geometrical form, orientation, position, and temporal extent of features	symbol with associated uncertainty measure	reaction to stimuli in the form of motor signal/vector or motivational state
Means of spatial registration	sensor coalignment	sensor coalignment, shared optics	geometrical transformations	spatial attributes of symbol, if necessary	geometrical transformations, if necessary
Degree of spatial registration	High	High	Medium	Low	Medium
Means of temporal registration	synchronization or estimation	synchronization	Synchronization	temporal attributes of symbol, if necessary	synchronization
Degree of temporal registration	High	Medium	Medium	Low	High
Fusion method	signal detection and estimation	image estimation or pixel attribute combination	geometrical and temporal correspondence, feature attribute combination	logical and statistical inference	weighted summation, fuzzy rules, subsumption, artificial potential fields, or neural nets
Improvement due to fusion	reduction in expected variance, improved detection	increase in performance of image processing tasks	reduced processing, richer, more accurate feature data	increase in truth or probability values	increase competence in real-time reactions

B. Line-of-Sight Sensors

1) Overview

LoS sensor proposed by [27] is comprised of an array of light sources and an array of light sensors. The light sources and sensors are assumed equal in number, equally spaced and placed along lines that are parallel to each other as shown in Figure 2. During one sensor cycle, the light sources are illuminated briefly and sequentially. Each source and each sensor is suitable for wide-angle emission and reception respectively and are modeled as points. The light sensors are synchronized with illumination from the light sources so every light sensor can detect the illumination of a light source when there is no obstruction between the light source-sensor pair. If a light sensor detects the illumination of a light source, then it is assumed no obstruction exists along the segment connecting the light source and sensor. If a light sensor does not detect illumination from a light source, this is evidence there is an obstruction between the light source and sensor.

Multiple obstructions could occur in collinear fashion along one source-sensor line. These can be detected and located, as explained later.

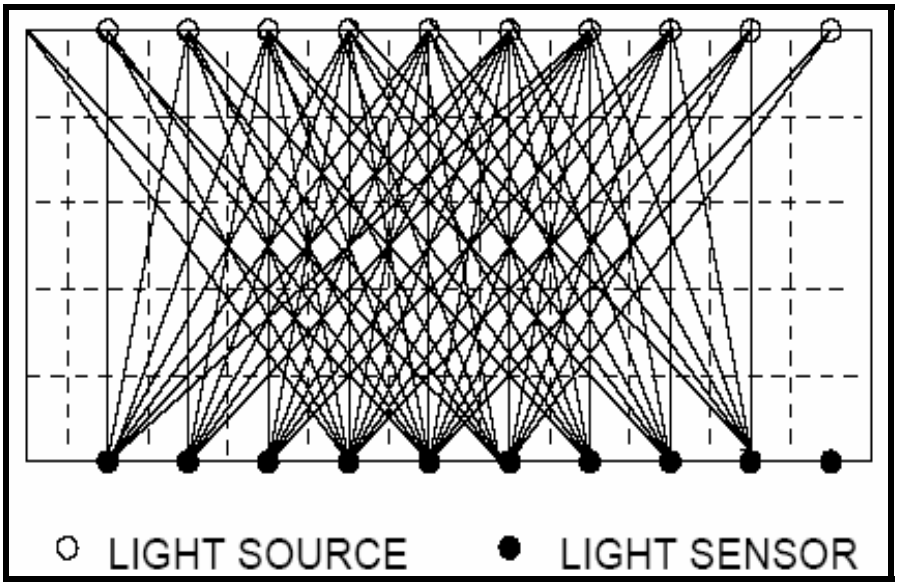


Figure 2. Diagram of LoS Sensor

The area between the linear arrays of light-sources and sensors is partitioned into squares, forming a two-dimensional grid [26]. The LoS sensor maintains a pair of two-dimensional data arrays with each element of the arrays map corresponding to a grid square. The first array holds evidence threshold values for each grid cell; the second array holds accumulated evidence of one sensor cycle. The evidence threshold value associated with a grid cell is the total number of light source-sensor segments passing through the grid cell. The values in the evidence accumulation array are the number of blocked light source-sensor segments intersecting the array element's corresponding grid square. The values of each element in the two arrays are compared to determine the percent of light source-sensor segments intersecting a blocked grid square. If the

percentage of blocked light source-sensor segments exceeds a given threshold, an obstruction is said to exist at the corresponding grid square.

2) Complexity

For a single LoS sensor, the number of light source-sensor segments is $S = L^2$, where L is the number of light sources and the number of light sensors. For each light source-sensor segment, the maximum number of increments made to the accumulation table is $S \times \max(H, W)$, where H is the height of accumulation array and W is the width of accumulation array. After all accumulations are complete, $H \times W$ comparisons are made between the accumulator array and the evidence threshold array followed by post-processing over a finite neighborhood. The computational complexity for the LoS sensor is then $O\left(S \times \max\left(H, W\right) + HW\right)$, with the maximum calculation occurring when all light source-sensor segments are blocked. The above complexity analysis assumes sequential computation, however, because L^2 lines are acquired in L clock pulses, speedup is obtained by concurrent computation. Tracking, however, can be done in two ways. When there is a limited number of moving people, using a list for each time interval facilitates nearest neighbor association. Thus, K people within the field of view require K^2 comparisons at most. However, if there are a large number of people present, with a maximum velocity V_{max}, there are $H \times W \times V_{max}^2$ steps to establish corresponding persons in two consecutive sets of video images. Conversely, given a pair of $N \times N$ video images, acquired from a pre-calibrated multi-camera system, $O(n^2)$ steps are required to locate the people, followed by $O(Kn)$ to establish correspondence. For each

correspondence, the depth computation takes a fixed amount of time, $O(1)$. Given the active sensing and distributed computation of the LoS sensor, multiple LoS sensor cycles can occur between each photo of the video cameras. Furthermore, if *K* hypothesis are generated by the LoS sensors, then the subsequent video-based validation/refinement requires processing over a finite area within each camera image, hence $O(K \times 1)$ steps over a fixed number of cameras.

3) Fusing with Camera Systems

Like computer vision systems, a Line-of-Sight (LoS) sensor is designed to determine location without requiring special badges or tags. However, the LoS sensor cannot identify a person directly, but can determine the location of numerous people within a Region of Interest (ROI). However, feature-level fusion of location data from the LoS sensor and image data greatly improves simultaneous tracking and identification of numerous people [11] [20]. A technique known as superresolution [28] has shown promise in extracting features from multiple images from a single camera provided the location of the object whose features are to be extracted is known. This type of information cannot otherwise be obtained through normal approaches "where only a few cameras are used to image the corridor from a few strategically selected locations. [4]"

Feature-level fusing camera images with the LoS sensor can be accomplished by using LoS collected evidence as a basis for the registration process. With this method, the LoS sensor collects evidence and creates a hypothesis for the location of detected people. The first camera in a round-robin ring of cameras compares this evidence, which maps to a global coordinate system, against cameras whose geometry within the focal

point of the camera could further refine the hypotheses. This method prevents unnecessary processing of pixels which the LoS sensor has already hypothesized a person does not exist, and also enables depth information to be applied to each pixel without further computation or information from other cameras. The refined hypotheses is then sent to the next camera in the round-robin ring until each camera observing the given region receives depth information and further refines the hypotheses. After the hypotheses have been refined by the final camera, the hypotheses are sent to a Sensor Data Processing Task (SDPT). By using real-time operating system methods [19], a LoS sensor can submit hypotheses that have not been refined by the cameras to the SDPT at significantly faster rates than the photo rates of the cameras. The SDPT can use the higher sample rates of the LoS sensor to maintain consistent labeling of people that appear to cross between camera photos.

To illustrate this method of fusion, Figure 3 on page 15 shows a computer graphics example of an image in which every pixel must be checked for detection, followed by registration of the pixels with other cameras to acquire depth information. Figure 4 on page 15 shows the same image with several of the pixels highlighted to pure black or white. Pixels in a deployed system would not be altered like the ones shown in Figure 4, but these highlights are demonstrate which pixels correspond to a person location hypothesis of the LoS sensor. The camera in Figure 4 only processes the highlighted pixels, with white highlights meaning the camera also detects a person, and black highlights meaning the camera does not detect a person at that pixel. The LoS sensor data contains depth information for the white highlighted pixels.

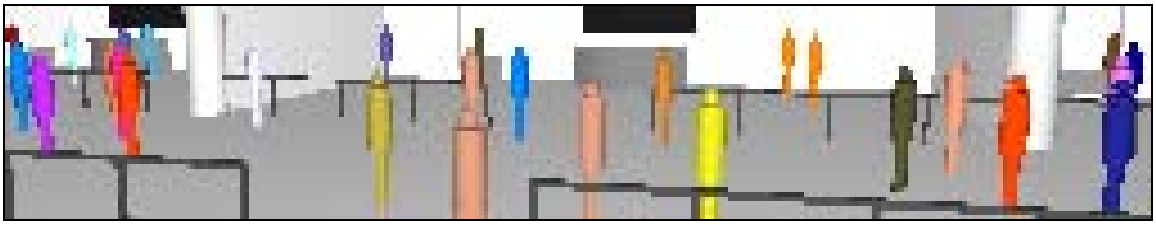


Figure 3. Example of an Image from a Camera for which Every Pixel Need Processed

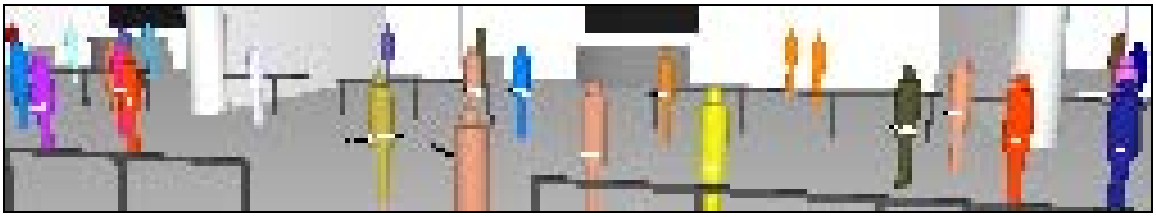


Figure 4. Example of an Image from a Camera for which Only Select Pixels Need Processed

For this type of fusion, the location hypotheses from the LoS sensor and the video cameras can be considered as a single logical sensor [19] from the perspective of higher processes. Each of the individual sensors is also available to higher level processes so information can be requested directly from the LoS sensor or video streamed through the network from one or more chosen cameras.

4) Deployment

Another significant challenge for a location aware system is deployment [14]. Location aware systems should be ubiquitous, unobtrusive, and inexpensive [26]. Distributing cameras over an area to locate people achieves the ubiquity and

unobtrusiveness desired. In addition, since cameras have decreased significantly in price over the last decade while improving in performance they can be considered inexpensive.

Location aware systems can cover wide areas containing a large number of people. For such systems, high communications costs make centralized algorithms undesirable [8]. Distributed algorithms operating over a sensor network provides the communications scalability at the cost of increased complexity [8]. The sensors of a location aware system should report location information with minimal latency; sensors should collaborate on processing data and submitting concise information to location aware applications [7]. This minimizes delays due to message transmission time as well as contention for and collisions of shared network media [7]. The recent expansion of wireless technology such as Wi-Fi and Bluetooth has improved the ability to network distributed sensors. However, inexpensive, high-bandwidth, low power, ubiquitous wireless coverage is still difficult to achieve [23]. For this reason, location aware computing systems using computer vision should perform data fusion and use location functions with distributed algorithms that minimize communications.

5) Scaling

A large-scale location aware system architecture proposed by [20] partitions the location aware system in a manner resembling a tree data structure. The location sensors of each service area are networked to a leaf server of the location aware system. Non-leaf servers direct leaf-to-leaf server communication and coordination, including hand-over of a person from one service area to another. Non-leaf servers distribute and direct user queries of the location system. A similar architecture has been proposed for an

acoustic sensor network, except that ROI membership of the sensors and organization or the servers is determined dynamically [7].

C. Line-of-Sight Sensor Operating with Camera System

1) Detection

Before determining the location of a person, they must be detected. Detection in computer vision is performed by comparing captured image data with expected image data and noting differences. Expected image data can be derived in several ways. One method uses an image of the scene known to contain no occupants. This method is useful if such an image can be captured and the scene, including lighting, does not change. A second method uses only the most recently captured image. This method considers changes to a scene that occurs during the time intervals between image captures and can detect a moving person, but cannot detect a person who stops moving. The person can again be detected if they resume moving. This permits self-calibration of a camera system and prevents false alarms due to slow changes over time such as lighting. Another method averages values of several of the most recently captured images. This method also detects people as they move, and continues to detect a person that stops until the stationary person appears in approximately as many images that were used to form the expected image. The person is again be detected by the system if they resume movement. This method also self-calibrates the camera system and preventing false alarms due to slow changes over time such as lighting. Detection in a LoS sensor is conducted in the same manner as computer vision, except the captured and expected data reflect accumulator array values rather than image data.

Image data from a camera results from a light rays traveling from points in the background scene, or person to be detected, to the focal point of the camera lens. The light rays measured by the camera are in the form of a cone whose tip is the focal point of the camera. The camera needs no equipment at the origin of the light rays for detection to occur. However, because the light rays all come to a single point, they are not well distributed in the area to be observed, and one person can conceal another person. In addition, if the view of one person partially obstructs the view of another person from the perspective of the camera such that they appear to overlap in the resulting image, a computer vision system may incorrectly classify the two distinct people as a single larger person. In a LoS sensor, the rays are segments with a light source on one end and a light sensor on the other. Because either a light source or sensor is required on each end of the optical segment, the LoS sensor is most effective in areas with opposing walls or other structures where lights or sensors can be attached. As the lights are both distributed and synchronized, the optical paths are more evenly dispersed over the area being monitored than they would be from a camera, making it more difficult for a person to be concealed by another person. Furthermore, because there is no single perspective from which the scene is viewed by the LoS sensor, it is more robust against a pair of people being incorrectly classified as a single person.

2) Location

Determining the location of a person using computer vision is a computationally complex task. When a pixel in an image indicates detection of a person, a spatial ray corresponds to the pixel's coverage area can be projected from the focal point of the

camera lens out to the scene [5]. The challenge is to determine the distance from the camera's focal point to the ray's termination point given every other pixel from the same image results in a similar ray also originating from the focal point [5]. Some systems this information assuming the bottom of a person is resting on the ground, and thus the distance to the person is estimated to be the intersection of the lowermost ray and the ground plane [2]. However, this technique is not effective if the bottom of a detected person is not resting on the ground, if the bottom of the person is not within the field of view of the camera, or if the view of the bottom of the person is obstructed. Another method observes features of a person while small perturbations are applied to the camera [16] [17]. Distance from the camera is estimated by the amount of movement in the features relative to the amount of movement of the camera. People further away appear to move less in the image than do people closer to the camera focal point. It is assumed the people being observed are not themselves moving. This will likely not be true in an area with a significant amount of pedestrian traffic.

Knowledge about people being observed such as size or known locations of observable features points on a person can also be used to determine the location of a person from an image. However, exploiting that knowledge requires identification of the person [5]. When people could be identified at points of entry to the observed area of the location aware system such as people scan identification tags upon entry to a controlled area. In this scenario, a computer vision system is provided identification and location hypotheses, which the system needs only verify, update, and maintain. However, a means of acquiring knowledge about a person, such as the previously mentioned entry

control system, may either not be available or may be bypassed by a given person. In addition, the track of a person within the observed area may be lost by the system. In these cases, there is no initial hypothesis of identity and location to verify.

Registration of multiple images containing correlated locations in the observed area from different positions improves the ability of a computer vision system to determine location and identity of observed people [5]. There are several techniques for registering information between two images, with the simplest being image rectification [5]. This technique can be applied when the area to be observed is flat and the geometry of the area is known [5]. Spatial rays can be projected from the focal point of the camera lens out into the scene, however the challenge is to determine the distance to the ray's termination point [5]. Adding multiple cameras adds additional sets of spatial rays projected from different focal points. For rectification of the multiple images, a plane is defined in the image space shared by the cameras. The intersection of the projected rays from each camera with the plane is calculated and value between these intersections are interpolated [5]. All areas of the plane that the rectified cameras have in the path of the presence of a person are hypothesized to be occupied by a person [5]. This technique does not provide a volumetric location of people in the image space, but only a hypothesis along the given plane.

In Figure 5, the white circles represent the location of two people in the scene to be located by a pair of cameras. The rectification plane contains the focal points of both cameras. The black circles of the figure indicate locations the rectification method would hypothesize contain people, even though these people do not exist in the scene. The

cameras then resolve the misperceptions by feature comparison, or else they track non-existent people.

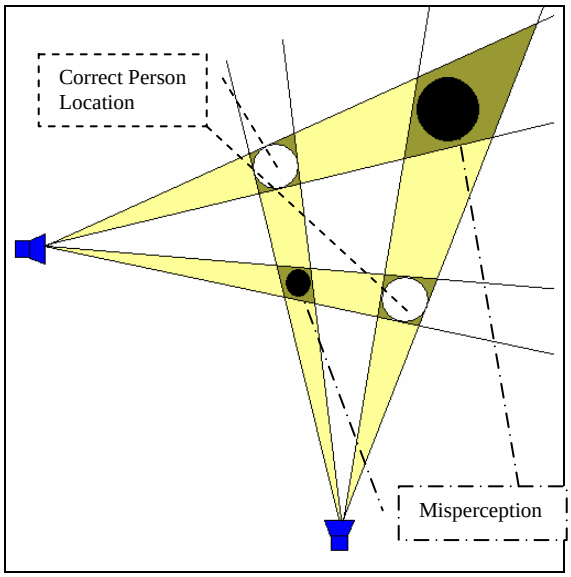


Figure 5. Two Camera Registration

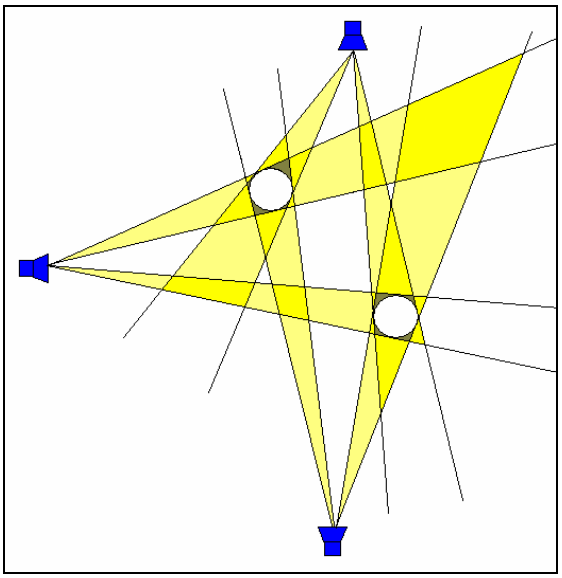


Figure 6. Three Camera Registration

Figure 6, is a replication of the scenario of Figure 5, except that a third camera is added whose focal point also lies in the plane of intersection. Information added by the third camera removes the two false hypotheses shown in Figure 5.

Including data from an increasing number of cameras can result in a location of the intersection plane transitioning from a hypothesized detection to a hypothesized non-detection. However, this cannot result in a transition from a hypothesized non-detection to a hypothesized detection. Increasing the number of cameras, then, can reduce the number of false detections. However, if there are any camera pixels with data that does

not differ from the background data, it would not indicate a detection and the corresponding region of the plane would be incorrectly assumed not to contain a person.

Because the accumulators of the LoS sensor map to grid squares of the area monitored by the sensor, data obtained from the LoS sensor is already a planar hypothesis of the locations of people. Consequently, no rectification of the data is required. In addition, because the LoS sensor observes blocked or non-blocked light, the performance of the system is not diminished by people whose appearance is similar to any background scene.

To register image data with LoS sensor data the image data can be rectified to the plane of the LoS sensor. After rectification, people are hypothesized to be at each location which the LoS sensor and the rectified image data both detect a person. An equivalent method of registering the LoS sensor data with image data is to register the LoS sensor data to the focal plane of the camera. This is done by assuming a projected line between each grid square of the LoS sensor and the focal point of the camera. The intersection of this line with the camera's focal plane is the corresponding location of the grid square in the image data, and is a rectification of the LoS grid square data to the camera's focal plane. As with the image registration technique, people are hypothesized to be at each location in which the rectified LoS sensor data and the image data both detect a person.

Rectifying the LoS sensor data to the focal plane of the camera offers several advantages over rectifying the image data to the LoS sensor. One advantage is the geometry of the area monitored by the LoS sensor is specified by the placement of the

light sources and sensors, whereas the geometry of the area monitored by a camera allows the focal point to be specified, but not the range of camera's detection. Consequently, the LoS sensor does not include data resulting from people outside of the ROI. Likewise, the LoS sensor only captures information from the plane in which it is installed. As a result, the LoS sensor does not include any information from people not part of the LoS sensor's plane. The absence of this extraneous information in the LoS sensor data results in less irrelevant information being sent through the network to the camera sensor, potentially reducing communication time.

In addition, the data from the LoS sensor can be interpreted by the receiving camera as location data. If there are no other cameras with which to coordinate data, the camera can do a pixel-level verification of the data from the LoS sensor to remove any points which the image data agrees can contain a person, without visiting every camera pixel in the process. The camera system may the use the remaining data as people locations and perform identification functions without incurring further network traffic costs for the location effort [26]. In this manner, the task of locating and identifying each person in the monitored area partitioned with the LoS sensor performing the majority of the location task, while the camera system performs the identification task.

3) Identification

For a camera system to identify a person, identifying features of a person must be detected [25]. Higher image resolutions result in more sampling of the image to facilitate identification of features [4]. Two means of improving resolution either replace the

camera with a higher resolution one, or combine multiple images of a person using a process known as superresolution.

In a camera system that does not use a LoS sensor, increasing the resolution of the camera increases the communication and computation time of the location and identification process. Extra communication time is incurred because data from each additional pixel must be transmitted to the other cameras [25]. Extra computation time is incurred because each additional pixel must be included in registration algorithm [25]. On the other hand, a camera system that includes the LoS sensor mitigates this cost by verifying pixels indicated by the LoS sensor data; a fraction of the camera's total pixels. By correcting LoS sensor data and using it as a basis for coordination and not sending additional pixel information between the cameras, the communication time between the cameras can be independent of the camera resolution. This, then, removes the need to tradeoff camera resolution and location time.

Video network systems have been built and tested to detect, identify, and track people [9]. These systems demonstrate the powerful capabilities of a video network system, but are still only capable of identifying an entity as being within a particular class of objects. An example system properly classifies detected objects at 99.5% for humans, 88.5% for groups of humans, 99.4% for vehicles and 64.5% for false alarms [9].

Superresolution can extract the features of a person. Superresolution operates by combining several images of a person in space and time, and aggregating the information into higher resolution images [27]. However, for superresolution to be successful, the location of the person must be known with respect to the camera for each image included

in the aggregation. Early results of this technique were "encouraging [4]." As with the higher resolution cameras, the camera system using superresolution can use the LoS sensor data to determine the location of people.

4) Tracking

Besides detecting, locating, and identifying people in a location aware system, the people must also be tracked over time. "The goal of practically all tracking algorithms is to obtain the probability density function (pdf) of the targets given all information," where a pdf is a range of continuous target location estimates [12]. This is a very complex task for a multi-sensor system tracking multiple people in a cluttered environment [16]. The Bayesian algorithm is a common statistical method for obtaining a pdf from multi-sensor data with correlated coordinates and time-stamps [6]. The Bayesian algorithm combines new sensor data with prior sensor data to maintain a population of hypotheses and a probability associated with each hypothesis, where hypotheses are a set of discrete target location estimates [6]. The Bayesian algorithm "is valid for non-Gaussian and nonlinear system/measurement models [24]." Each new sensor reading alters the probabilities of hypothesis and possibly adds new ones [6]. Use of the Bayesian algorithm is known as Multi-hypothesis tracking (MHT) and is considered a standard approach [22].

Because the number of hypotheses grows exponentially, a filter is used to prune hypothesis growth [6]. Many sensor fusion methods have their underpinnings in the Bayesian algorithm [6], such as the Probabilistic Data Association Filter (PDAF) [13]. The PDAF combines multiple hypotheses into a Gaussian mixture of allowable

associations [13], [24], thus reducing the number of individual hypothesis tracked [24]. The Kalman filter method is also derivative of the Bayesian algorithm [25] that assumes Gaussian distributions and linear models [25]. The Extended Kalman filter (EKF) is a recursive version of the Kalman filter, and is a class of Maximum Likelihood (ML) estimator algorithms [11]. Two other techniques that have been simulated to outperform implementations of the MHT, PDAF, and/or ML algorithms under certain conditions are the neural network approach simulated in [22] and the maximum a posteriori (MAP) algorithm [16]. The neural network and MAP approaches do not use behavioral estimations (velocity and/or acceleration for example) of the people being tracked which is required for many MHT and PDAF algorithms [16], [22].

Because the LoS sensor does not rely on ambient light for detection, the sensor can sample the observed area in significantly shorter time intervals than the 15 to 30 photos per second sample rate common in modern cameras. If the LoS sensor sample at a faster rate than the cameras, LoS sensor information alone can be used to alter the hypothesis probabilities between camera photos, thereby improving the ability of the location aware system to maintain track of the multiple people [25].

D. Chapter Summary

This chapter presents background for wide area surveillance using video cameras and LoS sensors. Important sub problems of location aware computing are introduced and existing approaches are described, along with the basic design of LoS sensor. The limitation of video-only networks is presented and the suitability of the LoS sensor for enhanced detecting, locating, and tracking multiple people is briefly explained.

III. Methodology

A. Chapter Overview

This chapter presents the methodology used to evaluate the Line-of-Sight (LoS) sensor as a location detection system for a video sensor network. First, the hypothesis and general approach to verify of the hypothesis is presented. The boundaries of the system to be simulated and evaluated are defined, including the services provided by the simulated system. Workloads submitted to the simulated sensor networks are explained and the levels of workload submitted to the system for this study is given, as well as the performance metrics used to evaluate the sensor configurations. Parameters of the system to be simulated, and parameters of the workload submitted to the system are also explained.

B. Hypothesis of Study

The purpose of the sensor network under consideration is to determine the location of all persons in the network's region of interest. The data from the sensor network is used to determine the identification and track of each person detected. The Line-of-Sight sensor is expected to provide a greater contribution to the location detection of a person than three networked cameras, and the Line-of-Sight sensor fused with any number of cameras is expected to offer fewer false alarms than three networked cameras.

C. Approach

To achieve the system's goals of locating individuals in a sensed area, a simulated layout is specified and divided into Regions of Interest (ROIs). One LoS sensor and three cameras are assigned to each Region of Interest (ROI) as components of a distributed sensor network. Each ROI is either a terminal with or without a desk, or a portion of a corridor that is part of an airport. The sensor networks fuse information from each contributing sensor to detect and locate individuals within their assigned ROI. Although a sensor may contain within its field of detection all or part of other ROIs, the data from a sensor is only used for detection in the ROI to which it is assigned. The effectiveness of the LoS sensor and each of the three cameras is determined through a full factorial simulation, in which all sixteen combinations of each sensor's data is either included or not included in a sensor fusion algorithm that determines the current location of people within each ROI. The hit rate, false alarm rate, and precision each sensor configuration is calculated and compared.

To meet the system goal of detecting the presence of people, models of people the sensors detect are also added to the simulation. To test the hypothesis that the LoS sensor provides a greater contribution to the location detection of a person than three networked cameras, and the LoS sensor fused with any number of cameras will offer fewer false alarms than three networked cameras, experiments with varying workloads are performed for each combination of LoS sensor and camera data inclusion. Performance metrics of each sensor configuration are compared.

D. System Boundaries

The system under test includes the cameras, LoS sensors, and the sensor fusion algorithm. The system and component under test is shown in Figure 7. The system includes the LoS sensor, networked video cameras, communication network, tracking algorithm, and system server. The system under test does not include the location aware application, building architecture, or people being tracked. The system under test also does not include the mechanism by which tracked people are initially identified. Another system provides an initial identification and location for each person as they enter the area covered by the sensor network. The system might be an entry control system using proximity badges and personal identification numbers. Thus, a security system or location aware application is considered a client or subscriber of the system under test and not a part of it.

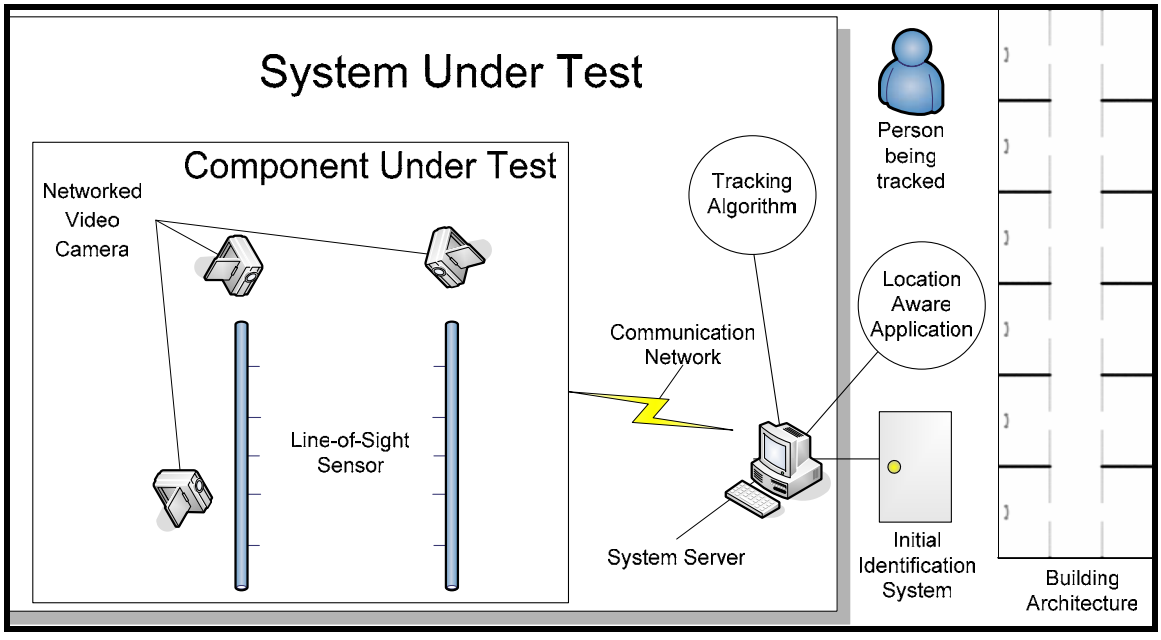


Figure 7. System and Component Under Test

The component under test includes the configuration of the sensors implemented in the system under test. Sixteen configurations are tested and compared. The configurations include every possible combination of one LoS sensor and three cameras.

E. System Services

The service provided by the system is the current location of each detected person within the system's area of coverage which is presented as a binary black and white image. Each pixel of the image corresponds to the center of a square. The squares represent three-inch-by-three-inch two dimensional grid areas of a global coordinate system. A white pixel in the image, or Boolean true, indicates the system hypothesizes the detection of a person at that grid location. A black pixel in the image, or Boolean false, indicates the system does not hypothesize the detection of a person at the corresponding grid location. This detection data provides the basis for the tracking algorithm which uses previous location detections to track individuals, as well as to access various static and dynamic characteristics from the camera images.

The possible outcomes of the system services is summarized in Table 2 on page 31. If the grid square containing center of the actual location of a person detected as containing a person, the system is considered to have correctly located that person. If the grid square containing center of the actual location of a person is detected as not containing a detected person, the system is considered to have not located that person. If the center of the actual location of more than one-person maps to a single component in the image, the component is considered to represent the correct location of only one

person, and the other people are considered to have not been located. Any components that do not contain the center of the actual location of a person constitutes a single false alarm of the system.

Table 2. System Outcome for Given Events

Event Name	Description	Outcome	Correctness
Person Location	One and only one person centroid maps to any of all pixels of a white component	White component reflects location of a person	Correct
α	Person centroid does not map to a white component	Location of person not given	Incorrect
Person Merging	More than one person maps to any of all pixels of a white component	White component only reflects location of one of the people	Correct for one person, incorrect for others
β	A white component containing no pixels mapped to by any person	White component reflects location of a person where no person is located	Incorrect

F. Workload

The workload submitted to the system under test is in the form of optical obstructions caused by people within the ROI, as well as clutter resulting from sensed people in adjacent ROIs. To simplify the pedestrian traffic model, a script derived from several random variables is used to simulate airport pedestrian traffic. The script gives the location of every person within the airport at 100 millisecond intervals for 100

seconds. The total number of points generated by the script is then 1000. Three types of people are simulated. The first type of person begins at a given terminal and remains within that terminal. The second type of person begins at a location within the corridor and travels along the corridor without entering any of the terminals. The third type of person travels from one terminal to another terminal. This person may begin or end either within a terminal or in the corridor en-route to a destination terminal.

Three of each kind of person is included per terminal. The number of terminals is twelve, so thirty-six of each type of person results in 108 people. Although the presence of a terminal contributes to the number of people in the model, the people are not assigned to specific terminals. Rather, the starting and ending location of a person is randomly distributed throughout the airport as is appropriate for the type of person. For example, although the twelve terminals contributes three stationary people each, three stationary people are not necessarily assigned to each terminal. Instead, each of the thirty-six stationary people are randomly assigned to the twelve terminals. The other two types of people are added in a similar manner.

To vary the workload submitted to the system, the number of people is set to 30, 60, or 90 people. For each test, the 30, 60, or 90 people are selected at random from the 108 total people. Each of the chosen people are randomly assigned a color from a list of 67 predefined colors.

G. Performance Metrics

The performance metrics measure the ability of the sixteen sensor configurations to locate people in terminals and corridors. The metrics include the hit rate, the

percentage of people uniquely located; the false alarm rate, as well as the precision of correct detections. The metrics for terminals with desks, terminals without desks, and corridors are computed separately. The percentage of correct detections is calculated by summing the number of detected people found within every terminal type or corridor and dividing the result by the total number of people present in every terminal type or corridor for a given simulation. The false alarm rate is the sum of all detections per simulation per area type that do not correspond to a person. Precision is the area corresponding to the number of pixels involved in the correct determination of location. Precision is measured for each correct location of a person and is averaged for each ROI. These metrics are gathered for each simulation and Analysis of Variance (ANOVA) is performed.

H. Parameters

1) System

The spatial diagram shown in Figure 8 is used for all simulations, reflecting a typical airport architecture. This layout represents an open airport corridor with twelve terminals. Figure 9 on page 34 shows the design of the corridor between terminals. Figure 10 on page 35 shows the layout of a terminal with a desk, while Figure 11 on page 35 shows the layout of a terminal without a desk.

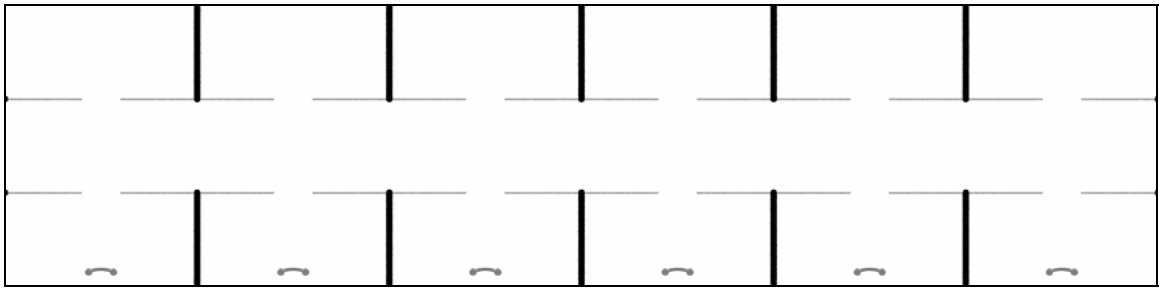


Figure 8. Spatial Configuration of Monitored Area used for Simulation

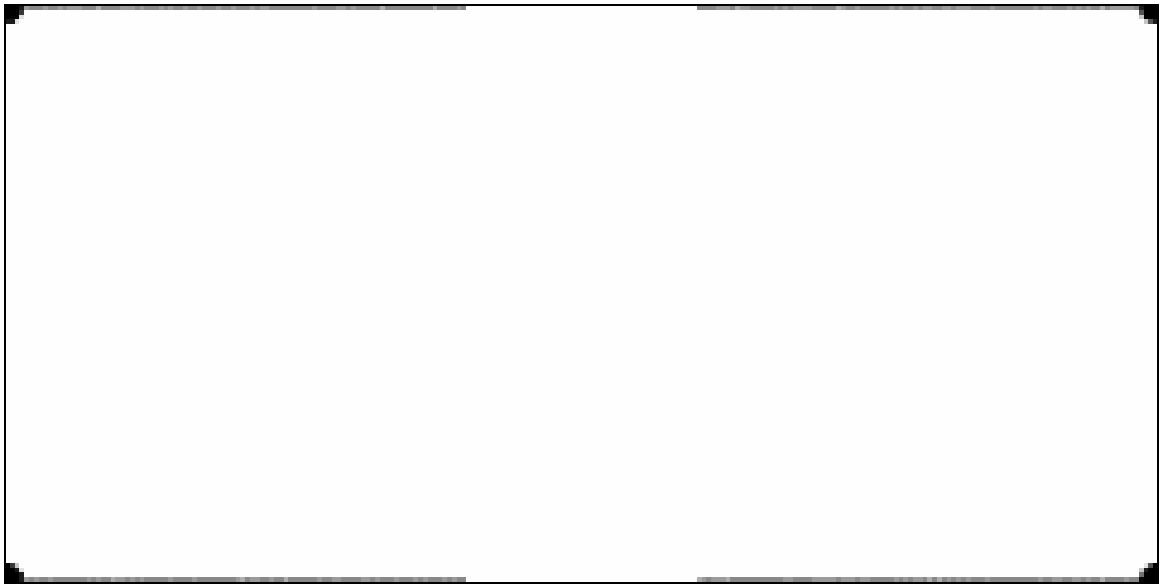


Figure 9. Corridor

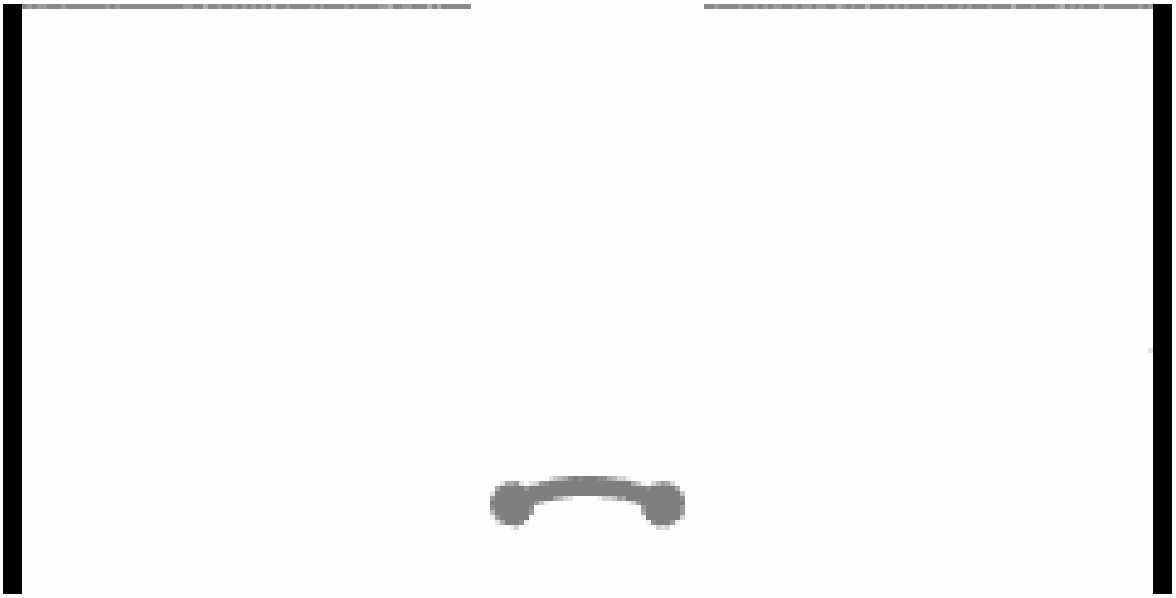


Figure 10. Terminal with Desk

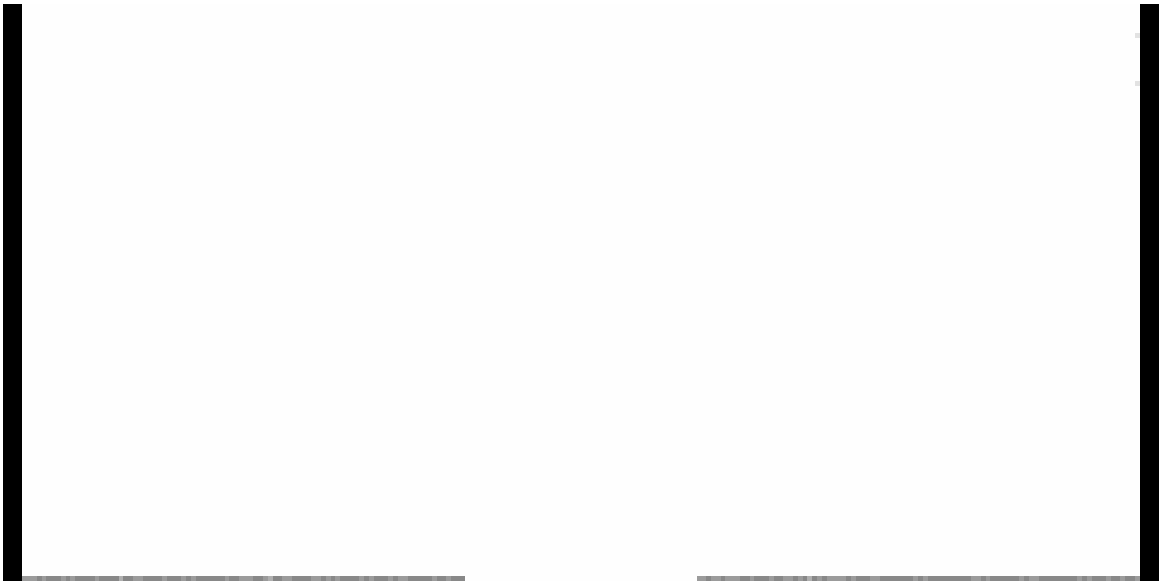


Figure 11. Terminal without Desk

Factors varied include the combination of cameras and LoS sensors used for detection. Sixteen camera and LoS sensor placement configurations of three cameras and one LoS sensor are used. Figure 12 on page 37 shows the placement of the cameras for the simulation. Figure 13 on page 37 shows the location of the three cameras sensing a terminal ROI without a desk. Figure 14 on page 38 shows the location of the three cameras sensing a terminal ROI with a desk. Figure 15 on page 38 shows the location of the three cameras sensing any portion of the corridor except the rightmost corridor ROI. Figure 16 on page 39 shows the location of the three cameras sensing the rightmost ROI of the corridor. Figure 17 on page 39 shows the placement of the LoS sensors. Each polygon represents a grouping of light rays that travel between the light sources of one side of the polygon to the light sensor of the opposite side of the polygon. In the corridor, the light sources and sensors are at the top and bottom of the polygons. In the terminals, the light sources and sensors are on the sides of the polygons. All cameras in both configurations are simulated to be placed 9.75 feet above the floor, whereas all LoS sensors are three feet above the floor. Each person is modeled as two coaxial cylinders topped by a sphere. The bottom cylinder represents the legs and is 0.9 feet in diameter, and the upper cylinder represents the body and is 1.155 feet in diameter. The sphere is one foot in diameter.

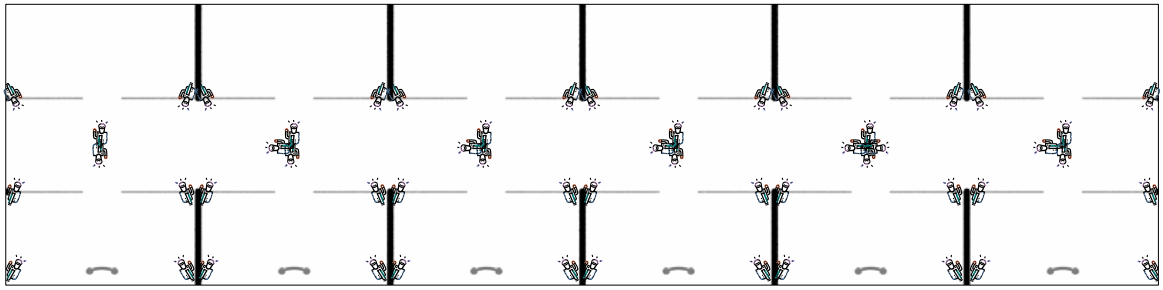


Figure 12. Camera Locations for Simulation

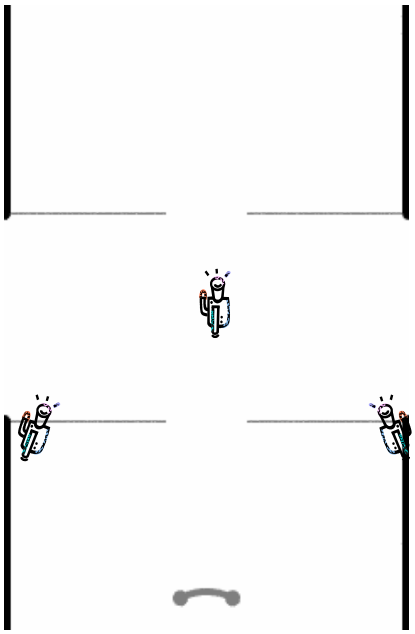


Figure 13. Camera Locations for Terminal ROI Without Desk

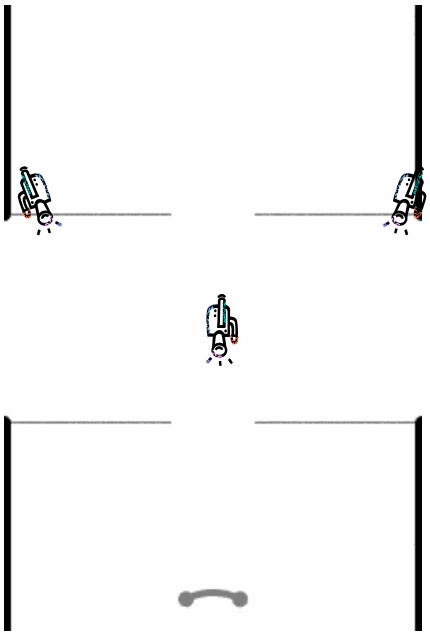


Figure 14. Camera Locations for Terminal ROI With Desk

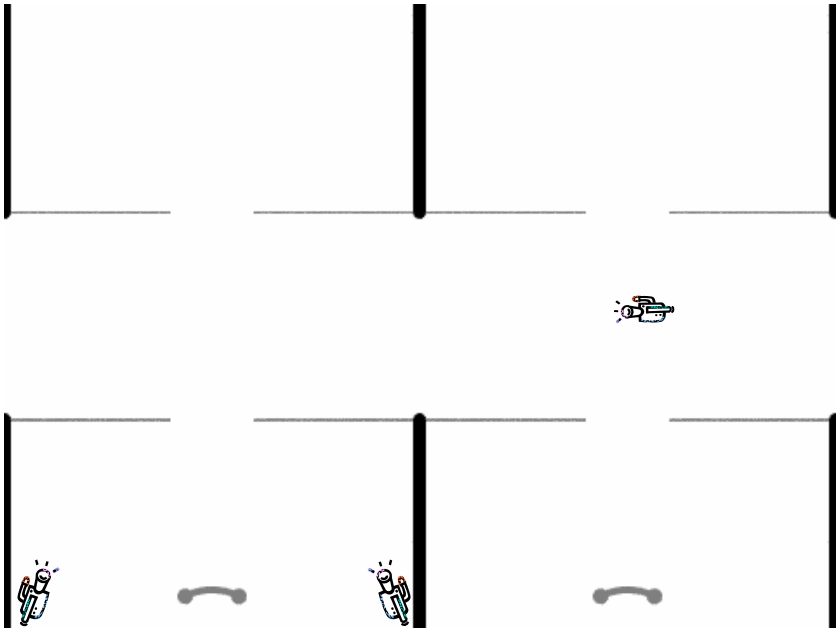


Figure 15. Camera Locations for Leftmost and Middle Corridor ROI

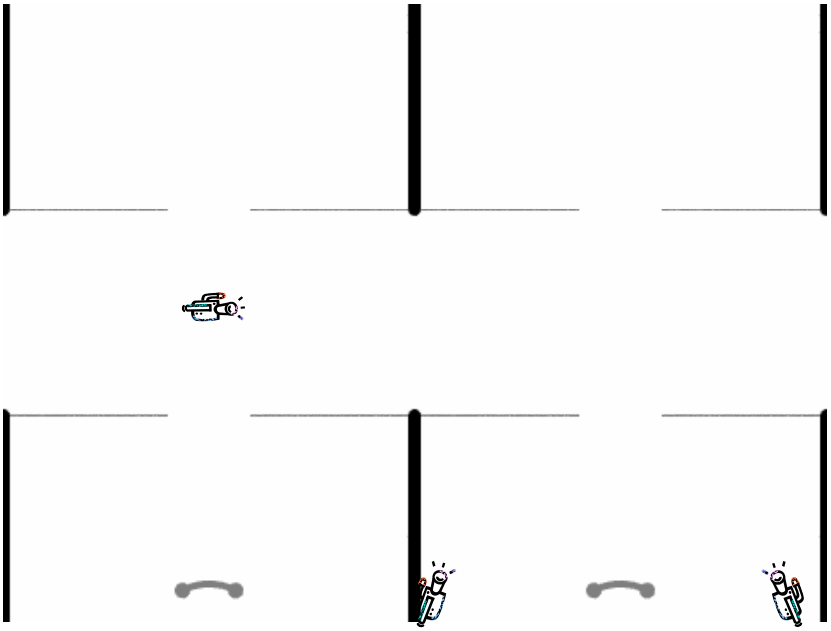


Figure 16. Camera Locations for Rightmost Corridor ROI

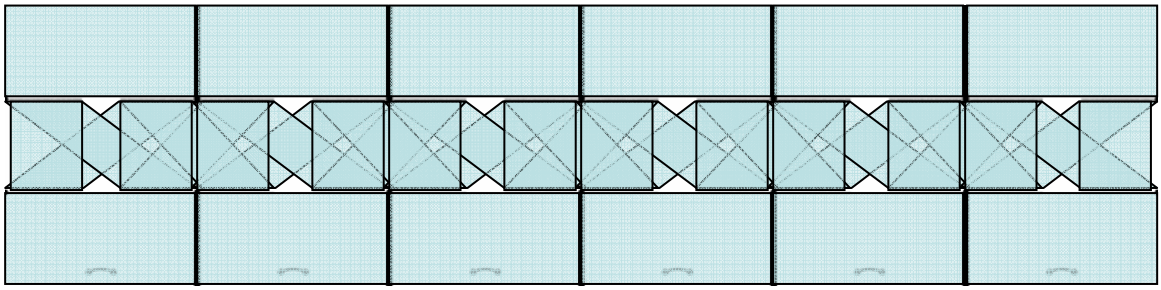


Figure 17. Line-of-Sight Sensor Placement for Simulation

2) Workload

The performance of the system is affected by the number of the people in the observed area. As the number of people increase, the probability of more than one person being perceived as a single larger person increases. This could lead to errors in

the location of one of the people. Differences in how the sixteen sensor configurations observe a scene may affect how well the system is able to distinguish people. A large number of people filling a scene may also increase the false alarm rate of a given sensor configuration.

I. Factors

The factors varied are the number of people being located and configuration of the sensors. Fifty simulations are run for three levels of people being located. Each scenario is replicated using each of the sixteen sensor configurations. The number of people in the area affect the performance of the system and can vary between the terminals and corridors. Three levels of population densities are 30, 60, and 90 people. Table 3 shows the factors and levels used.

Table 3. Factors and Levels for Simulation

Factor	Level		
LoS sensor	Present	Not Present	
Camera 1	Present	Not Present	
Camera 2	Present	Not Present	
Camera 3	Present	Not Present	
Number of People present in airport	30	60	90

J. Evaluation Technique

This system is evaluated using simulation. Because a prototype of the system has not been constructed, it cannot be measured. Camera and LoS sensor models are validated by presenting each simulation model with people at known locations and evaluating the resulting data from the sensors. Combinations of camera and LoS sensor models are validated in the same way. Current and historical detection data for every possible camera and LoS sensor combination is also visualized for the entire airport. Data from the simulation is validated by comparison with the visualization.

K. Experimental Design

Simulations are conducted with full factorial combinations of population levels using each sensor configuration, as shown in page 42, Table 4. With each run of the experiment testing every combination of area type and sensor configuration, a single simulation run conducts the equivalent of 48 individual experiments. The total number of runs to also vary the number of people is 3, which results in 144 experiments. The expected differences in the mean of the performance metrics for the sensor configurations is expected to be at least 0.0500, with a variance of up to 0.2000. If these assumptions are correct, at least forty-four replications are needed per experiment to distinguish statistically the performance of the sensor configurations at the 95% confidence level for a given scenario. For ease of statistical analysis, fifty replications are used.

Table 4. Simulation Factor Levels

LoS Sensor Present	Camera A Present	Camera B Present	Camera C Present	Area Type	Number of People
No	No	No	No	Corridor	30
Yes	Yes	Yes	Yes	Terminal, no desk	60
				Terminal, desk	90

L. Analysis and Interpretation

The hit rate gathered per test of the experiment represents a Bernoulli trial of whether or not each person is correctly located. The results of the Bernoulli trials are averaged as a mean probability representing the performance of the system for the hit rate. In addition, the mean number of false alarms per simulation for each area type is measured. Precision per simulation is found by averaging the area corresponding to the number of grid squares for each correctly located person. An analysis of variance is used determine the effect of individual factors and interrelationships between factors on each performance metric. For each relationship of factors whose ANOVA produces a p-factor less than 0.05, the mean and 95% confidence interval of the metric is determined, and a box-plot of the distribution is presented.

M. Summary

The methodology presented illustrates how the performance differences between the sixteen sensors configurations is acquired. The sensors are simulated as part of a distributed sensor network with given workloads. Identical simulations are run using

each of the sixteen configurations, and the performance metrics for each are compared. An analysis of variance is also conducted on the metrics to determine the effect of factors and factor interrelations on system performance.

IV. Simulation Results and Analysis

A. Chapter Overview

The purpose of this chapter is to present and analyze the results of the simulation study. First, the simulation model of the Line-of-Sight (LoS) sensor is described, followed by results of pilot studies used to determine parameters for the simulation model. A description of the camera model is also presented, and the method used to fuse the one or more modeled cameras with the LoS sensor. The results of the simulation are presented.

Two primary configurations are presented and analyzed. First, an Analysis of Variance (ANOVA) for all simulations the LoS sensor is used in is presented using hit rate, false alarm rate, and precision data gathered from the simulation. The mean value, 95% confidence interval, and box-plot graphs are presented for all factors for which the p-factor is less than 0.05. Second, ANOVA for all three-camera simulations but no LoS sensor is presented using all hit rate, false alarm rate, and precision data from the simulations. As with the configurations with the LoS, the mean value, 95% confidence interval and box-plot graphs are presented for all factors for which the p-factor is less than 0.05

B. Modeling the LoS Sensor

The LoS sensor is comprised of light sources and light sensors mounted on two opposing vertical surfaces. For this study, the light sources and sensors are assumed to be equal in number, equally spaced and placed along lines that are parallel to each other

at the same height. During one sensor cycle, the light sources are illuminated briefly and sequentially. The light sensors are synchronized with the illumination of the light sources so every light sensor can detect the illumination of a light source when there is no obstruction between the light source-sensor pair. If a light sensor detects the illumination of a light source, it is assumed no obstruction exists along the segment connecting the light source and sensor. If a light sensor does not detect the illumination of a light source, it is assumed that an obstruction does exist along the segment connecting the light source and sensor [27].

The area that contains the first and last light source and light sensor is partitioned into squares, forming a two-dimensional grid. The LoS sensor maintains a pair of two-dimensional arrays such that each element of the arrays maps to a corresponding grid square. The first array is called the comparison array, and the second array is called the accumulation array. The element values in the comparison array is the total number of light source-sensor segments intersecting the array element's corresponding grid square. The element values in the accumulation array is the number of light source-sensor segments that are blocked by an obstruction during a sensor cycle. The corresponding values in each element in the two arrays are compared pair-wise to determine the percentage of blocked light source-sensor segments. If the percentage of blocked light source-sensor segments exceeds a given threshold, an obstruction is hypothesized to exist at the corresponding grid square [26].

The simulation model of this system includes a segment array corresponding to the light source to sensor paths, the two-dimensional accumulation array, the two-

dimensional comparison array, and a lookup table associating each segment to intersecting grid squares. The LoS sensor simulation is based on the endpoints of two lines and the number of lights to be modeled. One of the lines defines the linear path the light sources are placed on, and the other line defines the linear path of the light sensors. The simulation model assumes the number of sensors is equal to the number of lights. Therefore, the number of segments is equal to the number of lights squared. The endpoints of the segments are equally spaced along and light source and light sensor lines, including the endpoints of the lines.

After the segments are initialized, the lookup table is created. The lookup table is generated by creating an array of squares whose geometry corresponds to the grid squares of the area. Each of the segments is tested to determine intersects with any of the four edges of each grid square. The results of the intersection tests are placed in the lookup table, which is maintained as a sparse matrix. It is assumed no segment changes state (i.e. goes from unblocked to blocked or blocked to unblocked) during a sensing cycle.

Obstructions are modeled as circles, each defined by center and radius. During simulation, all circles are compared with each segment of the LoS array for intersection. After all circle-to-segment intersection tests are complete, the segments that intersect with a circle are compared with the lookup table and the results placed in the accumulation array.

C. Pilot Studies

A pilot study was conducted to evaluate the effect of various parameters of an early simulation model of the LoS sensor. The object detection threshold for the LoS

sensor was 90%. A later study found that a higher threshold, 97.5%, offers better LoS sensor performance. The fitness criterion chosen for the study is the probability an object exists in a location when the sensor indicates that it does, or P(Object|Object Sensed), and the probability that an object does not exist in a location when the sensor indicates that it does not, P(No Object|No Object Sensed). These probabilities are measured through many simulations of the LoS sensor simulation model detecting uniform random distributions and of random numbers of objects. The simulations were performed with the number of lights in the sensor, the length to width ratio, and resolution of the sensor's output varied. Table 5 shows the levels for each of these factors.

Table 5. Factors and Levels for LoS Sensor Pilot Study

Factor	**Level**			
Relative Resolution	1	4	8	
Height/Width Ratio of Sensed Area	0.5	1	2	
Number of Lights/Sensors	4	16	32	64

Relative resolution refers to the size of a sensor grid square compared to a unit grid square. For example, a relative resolution of one means that a LoS sensor grid square exactly corresponds to a unit square of the simulation. A relative resolution of four means that the LoS sensor divides a unit square into 16 sensor grid squares in a grid, four high and four wide.

The height/width ratio refers to the dimensions of the LoS sensor. For this factor, the length of each line of light sources and sensors remained the same, but the distance between them was varied. For example, a height/width ratio of 0.5 spaces the light sources equally along a ten-unit length line, and the sensors along a parallel line with the

two parallel lines spaced five units apart. For the height/width ratio of 2, the two parallel lines of light sources and sensors would be placed 20 units apart.

The number of light sources/sensors are the number of light sources that are evenly distributed along one parallel line. An equal number of light sensors are evenly distributed along a parallel line.

For each combination of factors, 30,000 simulations were conducted and evaluated for a total of 1,080,000 simulations. The product of P(Object|Object Sensed) and P(No Object|No Object Sensed) is used to assign empirical performance values for each combination of factors. The combination of factors that produced the best performance using this method was found to be a LoS sensor with relative resolution of 1, a height/width ratio of 1, and 16 light sources and sensors. Analysis of variance was not conducted for the pilot study. Figure 18, Figure 19, and Figure 21 present partial results which summarize the effect of the parameters for this study.

Figure 18 shows that increasing the resolution of the LoS sensor increases the probability that the sensor model will correctly identify an empty grid square, but decreases the probability it will correctly find an occupied one. The reasons for these trends is likely because as the resolution increases, the number of grid squares between the lights also increases. As the resolution increases to 4, 16 lights span a length of 40 LoS sensor grid squares. This may result in many grid squares with no light rays intersecting them, which is effectively observed by the LoS sensor. As the resolution increases to 80 grid squares between the lights, P(Object|Object Sensed) becomes less than 0.50. It is likely at this resolution many grid squares existed in the simulation with

only one LoS sensor light ray intersecting it. For these grid squares, any object existing along the line of the light ray will result in the grid square hypothesizing it contains an object. A large number of grid squares intersected by a single ray will result in an increased number of false alarms.

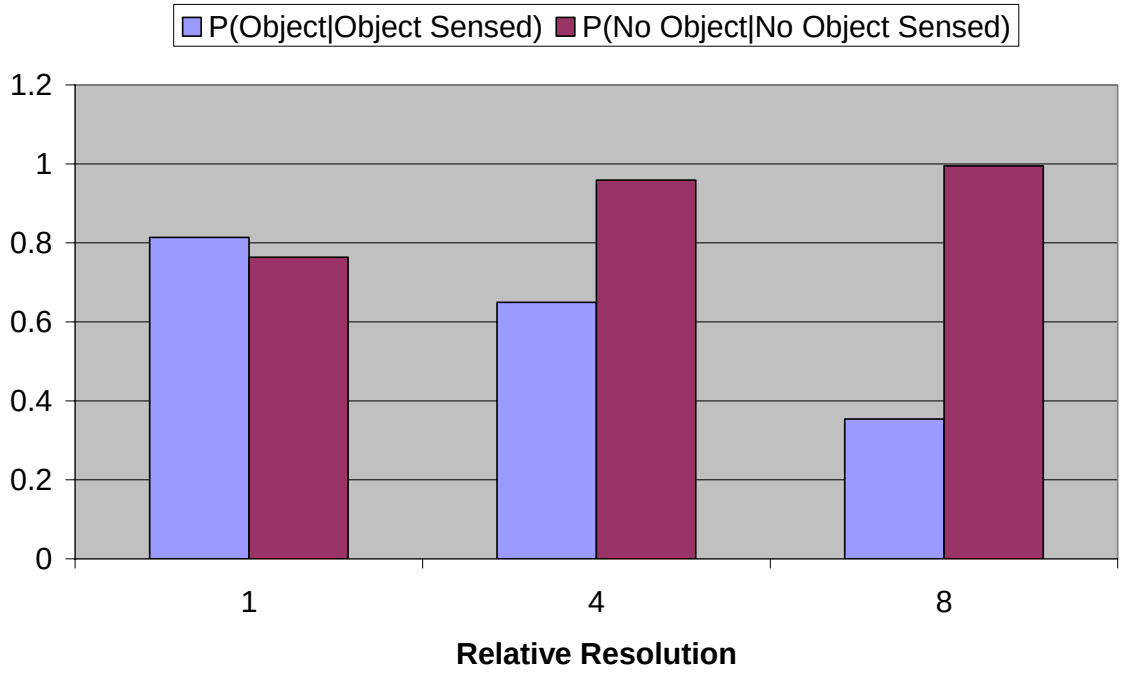


Figure 18. Pilot Study Results for Relative Resolution with Height/Width Ratio of 1 and 16 Lights/Sensors

As the resolution of the LoS sensor approaches infinity, the sensor performance degenerates such that most grid squares will not be intersected by any light rays, and

most of the intersected grid squares will be intersected by only one light ray. The sensor would perform as though the accumulator and comparison arrays were not present.

Figure 19 shows that increasing the distance between a pair of parallel light and sensor lines increases the probability of correctly determining the grid square occupied, but decreases the probability of correctly determining a grid square is unoccupied. The reason for this is shown in Figure 20. Figure 20 shows a black circle detected by a LoS sensor. The diagonal lines are tangent lines between the edge of the sensor and the detected object. The gray areas are object areas that the light cannot pass through, even though the areas are not occupied by the detected object. The result is the object appears to stretch towards the sensor lights and sensors. Increasing the height to width ratio of the sensor increases this stretching effect. The result of this stretching is an increase in area incorrectly showing occupation by an object, as was seen in Figure 19.

Figure 21 shows that the probability of correctly identifying that a grid square contains an object increases as lights increase from four to sixteen, and declines as more lights are introduced. Because the relative resolution is one, and the height to width ratio is also one, the LoS of Figure 21 is a ten by ten grid of squares. It is likely the poor performance of the four light source/sensor level is due to having more than two grid squares between each light. This results in several undetectable areas for the LoS sensor. It could also be the threshold for the LoS of this study was too low. Since decreasing the threshold level of the LoS sensor increases the number of false alarms, it is possible that increasing the number of lights also increased the false alarms that were not properly removed by application of a threshold.

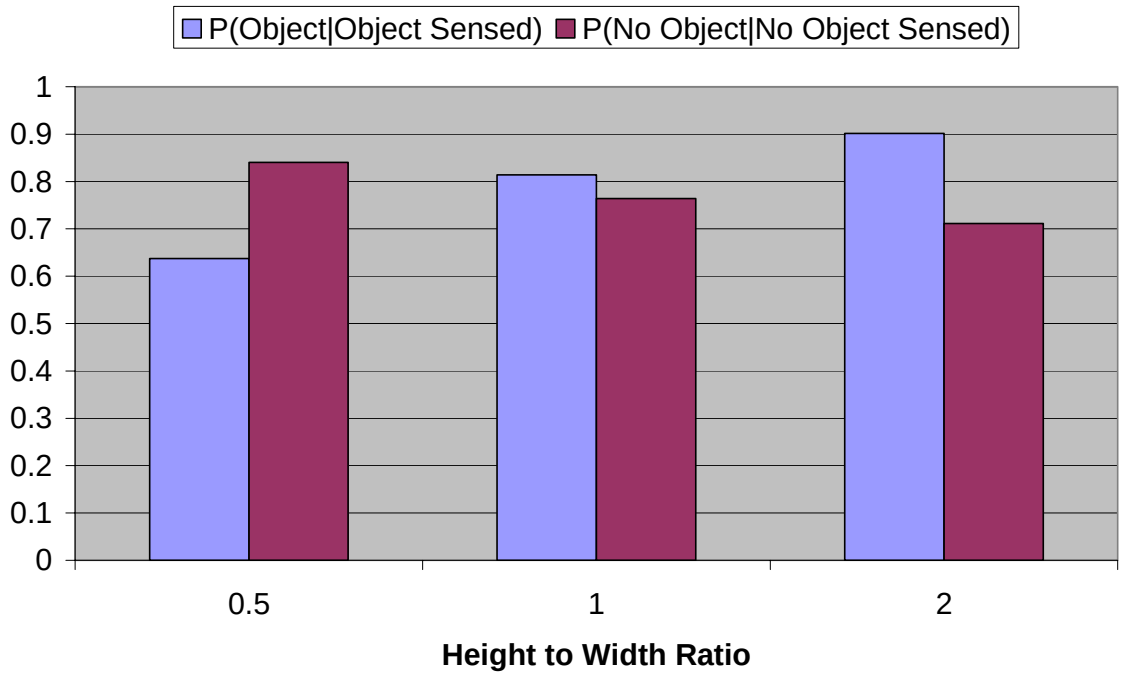


Figure 19. Pilot Study Results for Height to Width Ratio with Relative Resolution of 1 and 16 Lights/Sensors

The results of this pilot study indicate the desired resolution of the grid squares be considered when deciding the number of lights and sensors. The simulation designed for this project uses a height and width of a grid square was three inches. This size grid square results in each region of interest of the simulation being represented by a 120×240-pixel image. The diameter of each person is 13.86 inches so using three-inch grid squares ensures at least four grid squares in diameter. With a four grid square diameter, Gaussian-smoothing algorithms may be performed on a LoS sensor image without losing detection of the person.

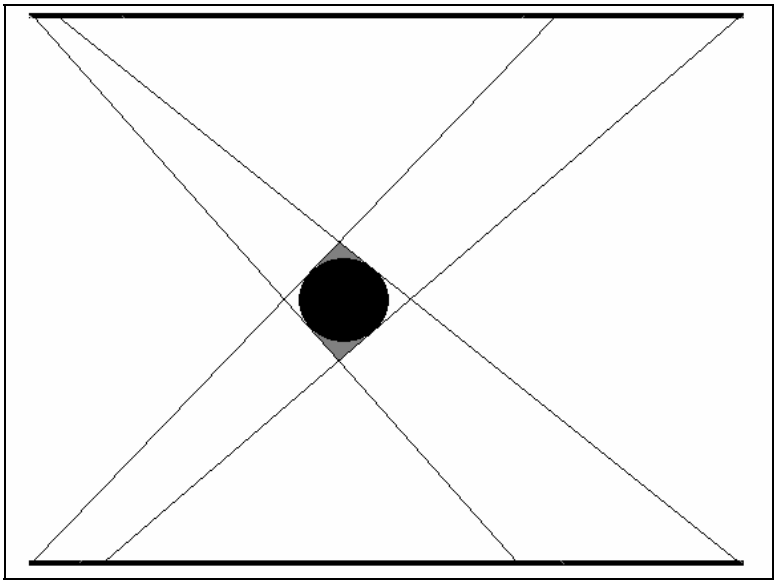


Figure 20. Example of Los Sensor Detecting Object

Two configurations of LoS sensor are used for this simulation. Because the terminals offer two walls without obstructions such as an entryway, one sensor is placed on each of these two walls. In addition, the pilot study indicated a better positive detection rate might be achieved by a separation of lights from sensors, while the area of each detected object may consequently increase. The one sensor provides sensing capability for the majority of the terminal. Figure 22 on page 54 shows the comparison array of a LoS sensor in a terminal without a desk. The 45 light sources and sensors are placed eight inches apart on the left and right edges of the figure. Each square of the image corresponds to a grid square observed by the LoS sensor, and the brightness of each square represents the number of light source to sensor segments intersecting it. The square in the middle of Figure 22 is the brightest because it has the most segments intersecting it. The top and bottom squares of Figure 22 are darker fewer segments

intersect them. Figure 23 on page 54 is the same LoS sensor shown in Figure 22, except each square is white if any segment intersects it and black if no segment intersects it. This figure reflects the coverage area of the same LoS sensor shown in Figure 22. Figure 24 on page 55 shows a comparison array of a similar LoS sensor in a terminal with a desk, and Figure 25 reflects the coverage area. Although no detection can occur in sections along the left and right edges of the terminals, the area of theses sections is less than the area of the people to be detected.

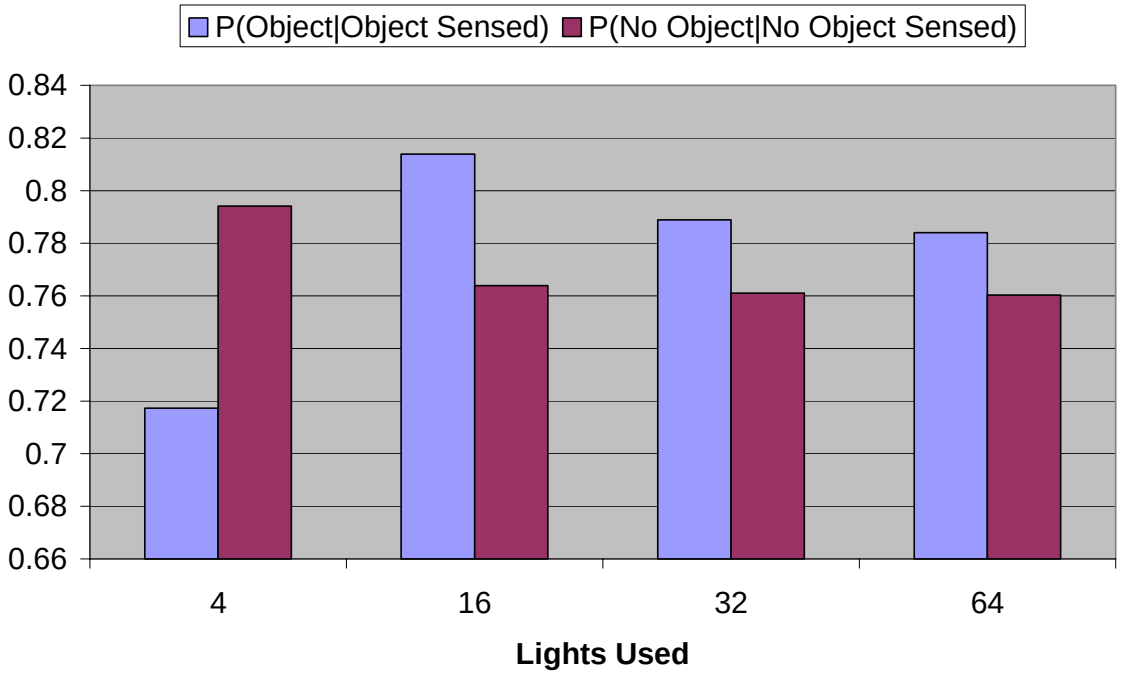


Figure 21. Pilot Study Results for Number of Lights/Sensors with Relative Resolution of 1 and A Height/Width Ratio of 1

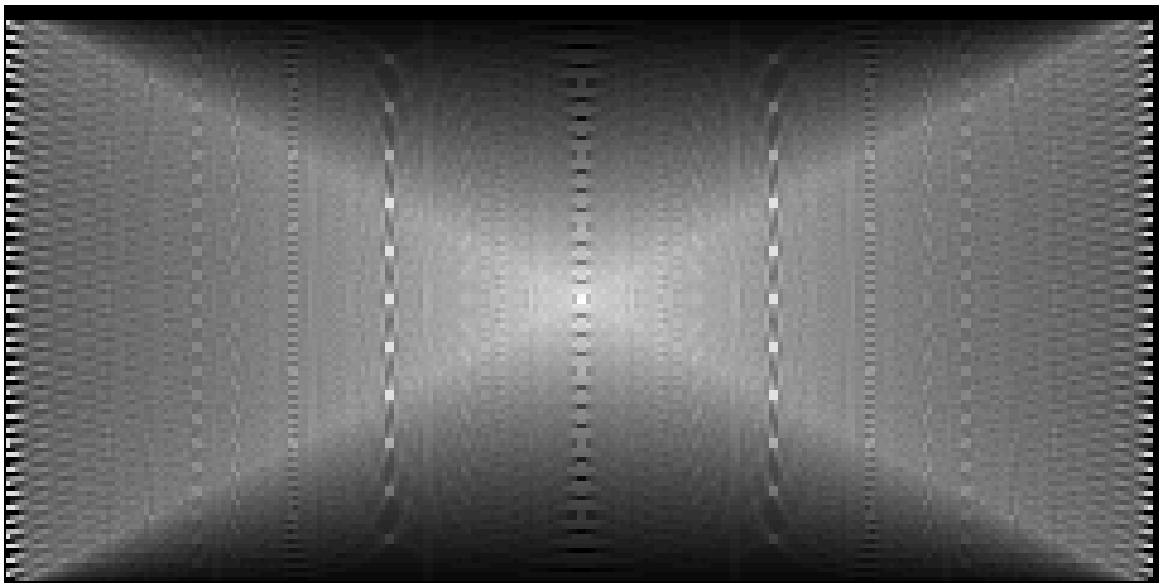


Figure 22. Comparison Array in Terminal without Desk

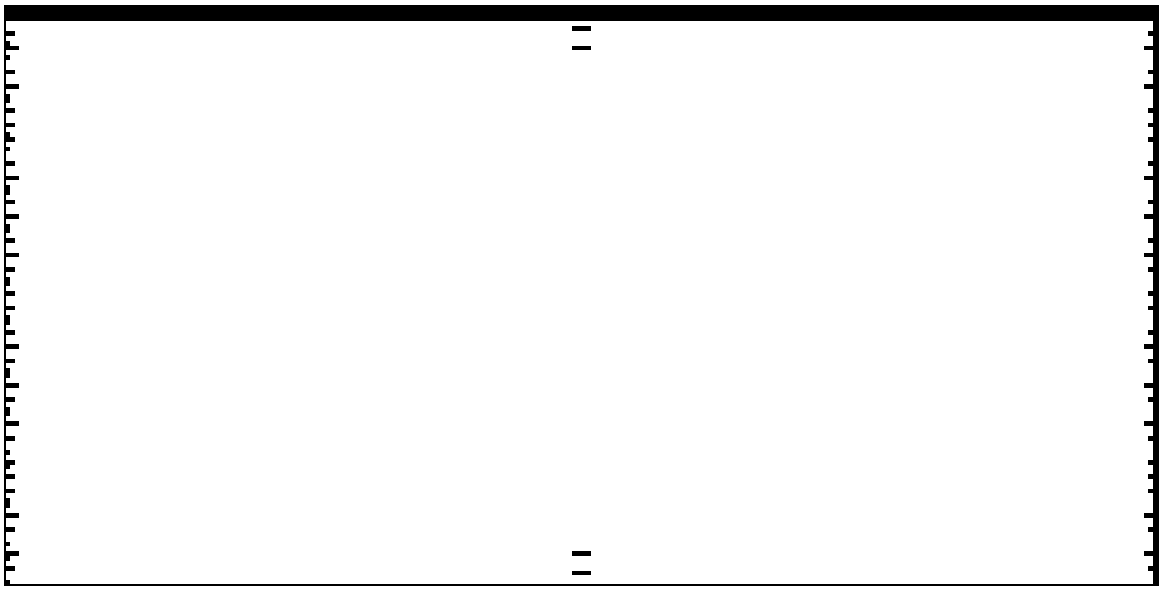


Figure 23. Coverage of LoS Sensor in Terminal Without Desk

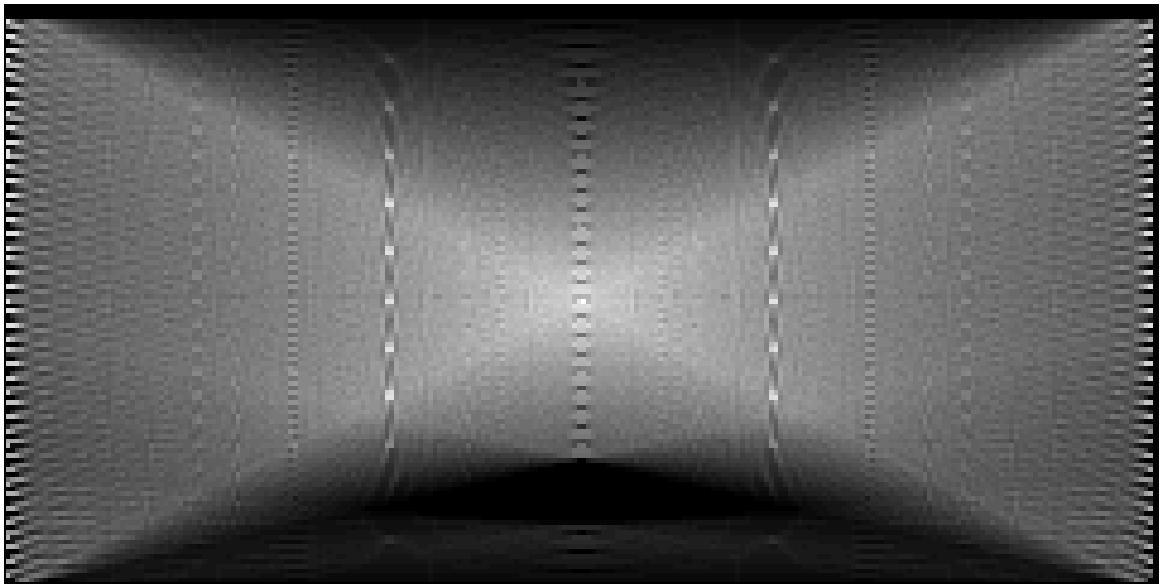


Figure 24. Comparison Array of LoS Sensor in Terminal With Desk

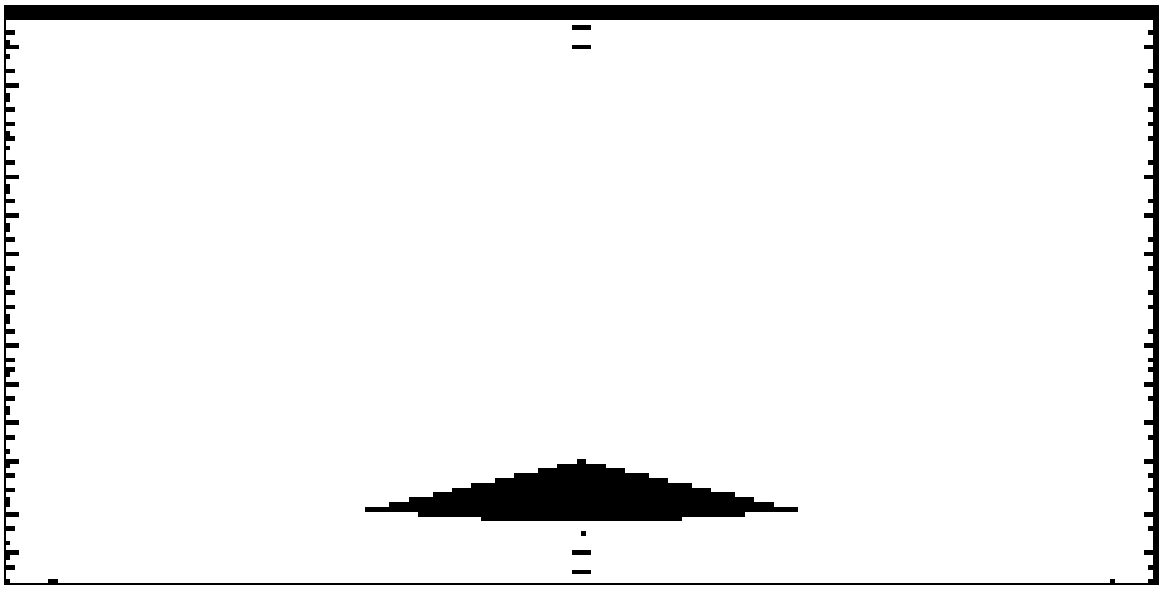


Figure 25. Coverage of LoS Sensor in Terminal With Desk

The corridor does not provide a pair of opposing surfaces from which the entire corridor can be monitored. Instead, two LoS sensors are used for each corridor ROI, one on either side of the opposing entryways. To cover the area between the opposing entryways, the LoS sensors synchronize and observe each other's light sources. The LoS sensors in the corridor observe the light sources in the nearest LoS sensor of the adjacent corridor LoS. The result of combining the sensing capabilities of the LoS sensors provides sensing capability for the majority of the corridor with the exception of areas directly in front of the entryways. Figure 26 on page 58 shows the combined comparison arrays of the LoS sensors in the middle of the corridor. There are two sets of 16 light sources and sensors placed eight inches apart on the top and bottom edges of the figure. As with the LoS sensors in the terminals, each square of the image corresponds to a grid square observed by the LoS sensor, and the brightness of each square represents the number of light source to sensor segments intersecting it. Figure 27 on page 58 is the same LoS sensor as shown in Figure 26, except each square is white if any segment intersects it and black if no segment intersects it. No detection can occur in the triangular areas between the sets of light sources and sensors.

The leftmost LoS sensor does not have a corresponding sensor further to the left to form additional light source-sensor pairs. Figure 28 on page 59 shows the comparison array of these LoS sensors. Figure 29 on page 59 shows the corresponding coverage. Note that the coverage of the sensor declines at the leftmost extremity of the sensed area.

The rightmost LoS sensor also does not have a corresponding sensor further to the right with which it can form additional light source-sensor pairs. Figure 30 on page 60

shows the comparison array of these LoS sensors. Figure 31 on page 60 shows the corresponding coverage. Note that the coverage of the sensor declines at the rightmost extremity of the sensed area.

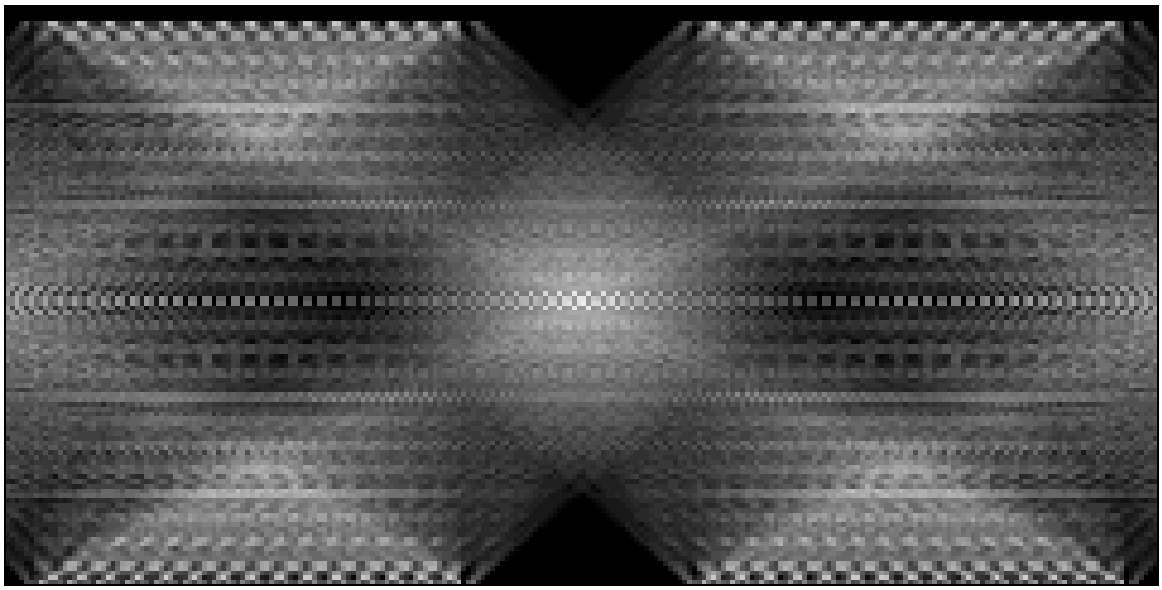


Figure 26. Comparison Array of LoS Sensor in Corridor

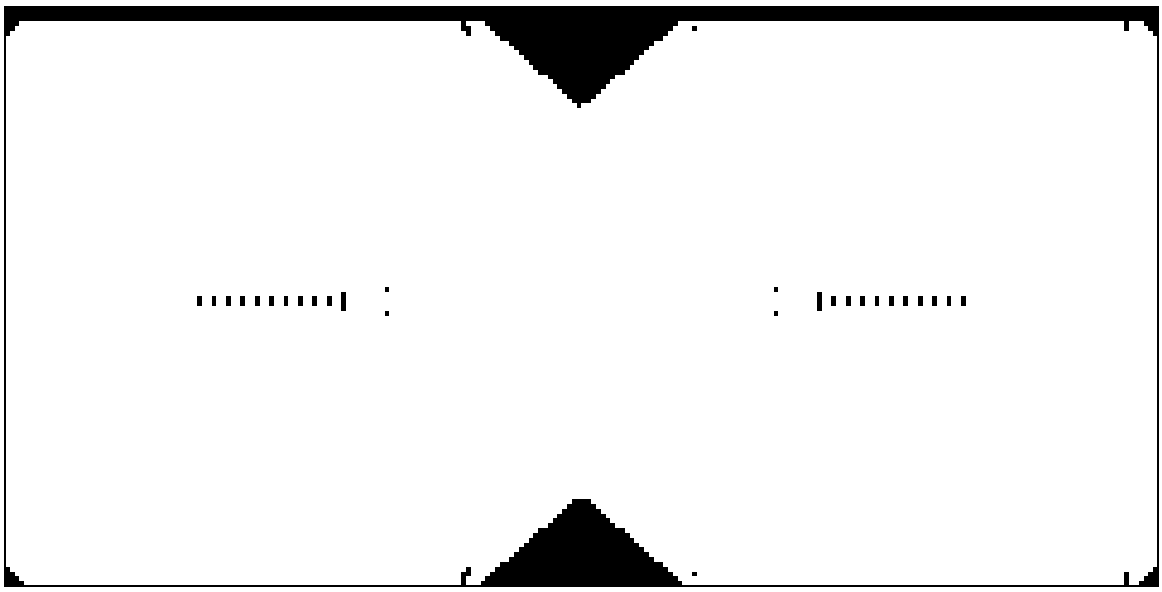


Figure 27. Coverage of LoS Sensor in Corridor

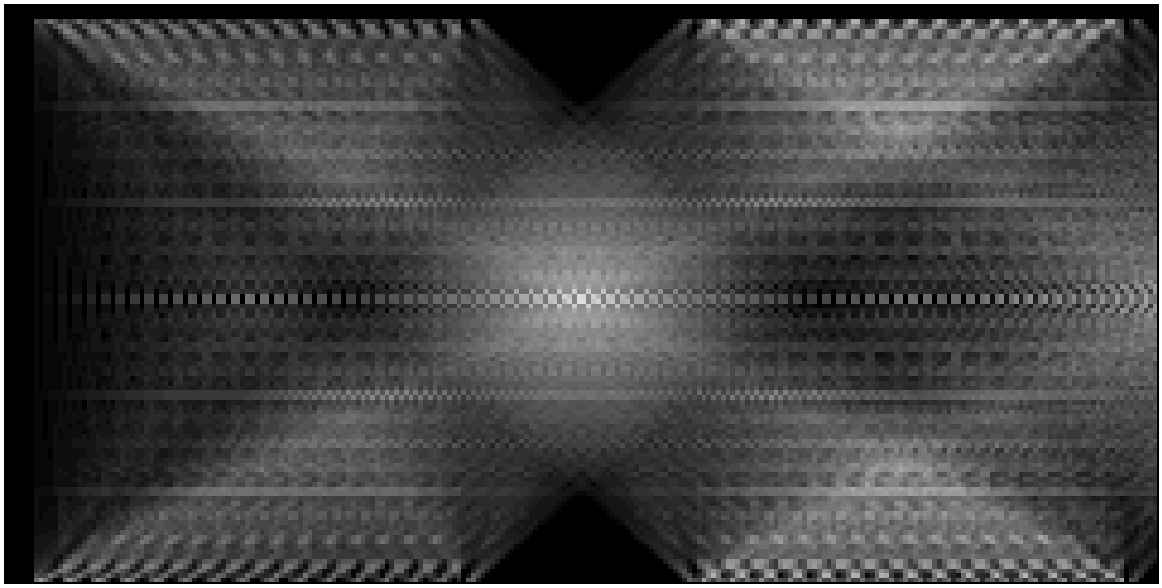


Figure 28. Comparison Array of Leftmost Corridor LoS Sensor

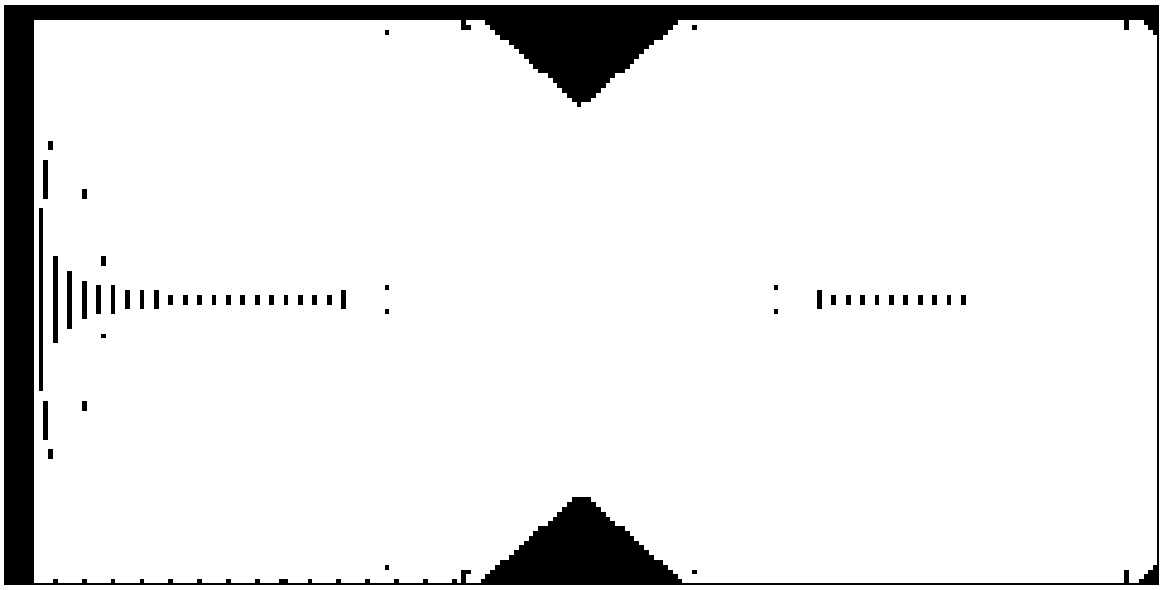


Figure 29. Coverage of Leftmost Corridor LoS Sensor

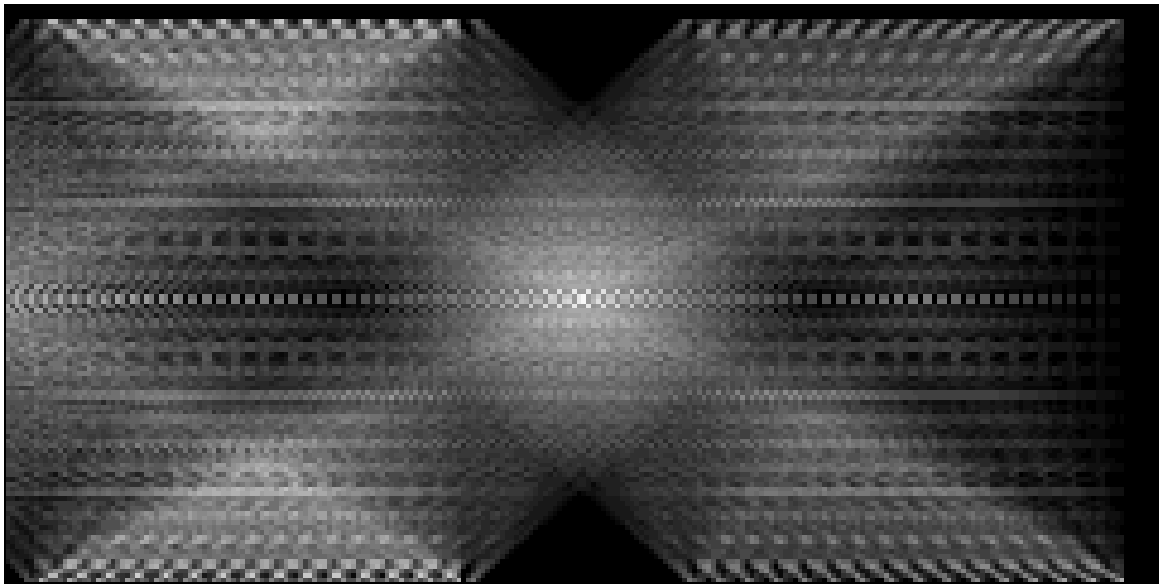


Figure 30. Comparison Array of Rightmost Corridor LoS Sensor

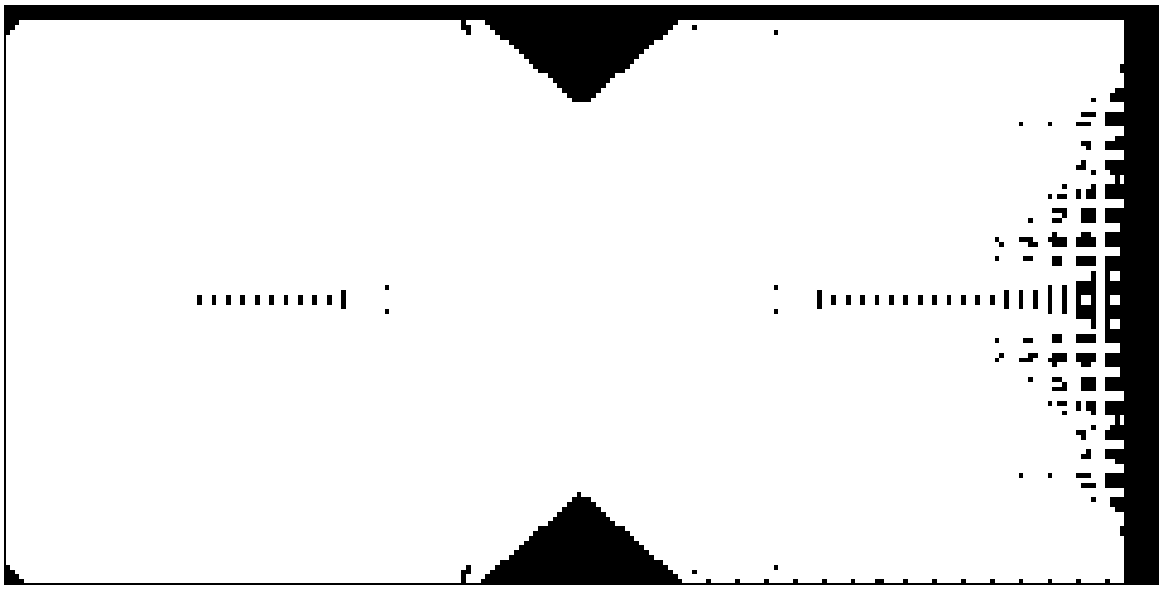


Figure 31. Coverage of Rightmost Corridor LoS Sensor

D. Modeling the Camera

A computer graphics program POV-Ray (version 3.6.1) is used to model cameras, and produce video images for the simulation. POV-Ray is a ray-tracing graphics program, that calculates values for each pixel individually which is desirable to determine pixel-level accuracy of simulated images. A three-dimensional model of the airport to be monitored is defined in POV-Ray, including a model of each simulated person to be located. The model of the airport is divided into 18 ROIs with three cameras each, totaling 54 cameras. Each of the 54 corresponding images are made by setting the POV-Ray's camera variables to a particular perspective and having each camera record an image of the scene simultaneously.

The airport has as six terminals on each side of a corridor. The terminals on one side of the corridor contain a desk whereas the terminals on the opposite side of the corridor do not. The corners of the desk are cylindrical such that the desk can be approximately modeled as circles for the LoS sensor. The colors of the airport are limited to colors with equal red, green and blue (RGB) components such as white, black, and gray to distinguish it from the people limited to colors with non-equal RGB components.

Each person is modeled with a cylinder representing both legs, one cylinder representing a body, and one sphere representing a head. The cylindrical shape is used because the LoS sensor model's obstructions as circles. The center and radius of a body is equal to the corresponding obstruction circle center and radius of the LoS sensor to ensure correspondence between the two types of sensors. Each person is also assigned a

single solid color. The color of each person is chosen at random from 67 pre-defined colors provided by the POV-Ray software package. The 67 colors used do not include colors with equal RGB components such as white, black, and gray.

To initialize the cameras, an image without people is captured at each camera location. This image is a background or static image of the scene observed by each camera. Figure 32 on page 63 shows three camera background images for a terminal with a desk. Figure 33 and Figure 34 show background images for a terminal without a desk and background images for a corridor respectively.

During simulation, each image captured is compared to its corresponding background image. Pixels equal in value to the background image are changed to black; pixels that differ from the background image are changed to white. The resulting image is rectified to a plane in the camera's corresponding ROI. The plane is three feet above and parallel to the ground plane, and is confined by the region defining the ROI. Each pixel of the rectified image corresponds to a grid square of the ROI. A white pixel of this image indicates the camera detects an obstruction at that grid square; a black pixel indicates an obstruction is not detected at that grid square. Figure 35 on page 64 shows three camera images for a terminal without a desk. The top three images were generated by POV-Ray. The next three images show the difference between the generated images and the background images. The final image is the result of rectifying the image to the ROI plane. Figure 36 on page 65 shows the same process for three camera images for a terminal with a desk, and Figure 37 on page 66 repeats the process for three images for the corridor.

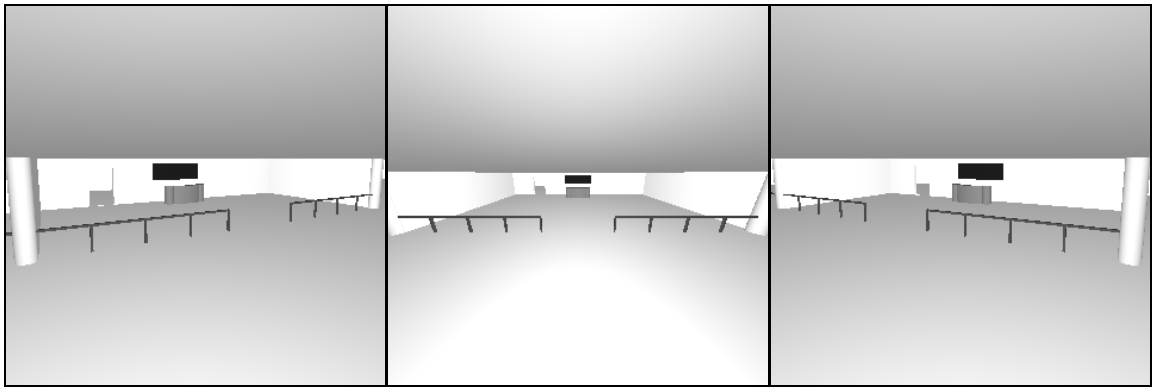


Figure 32. Three Camera View of Terminal with Desk

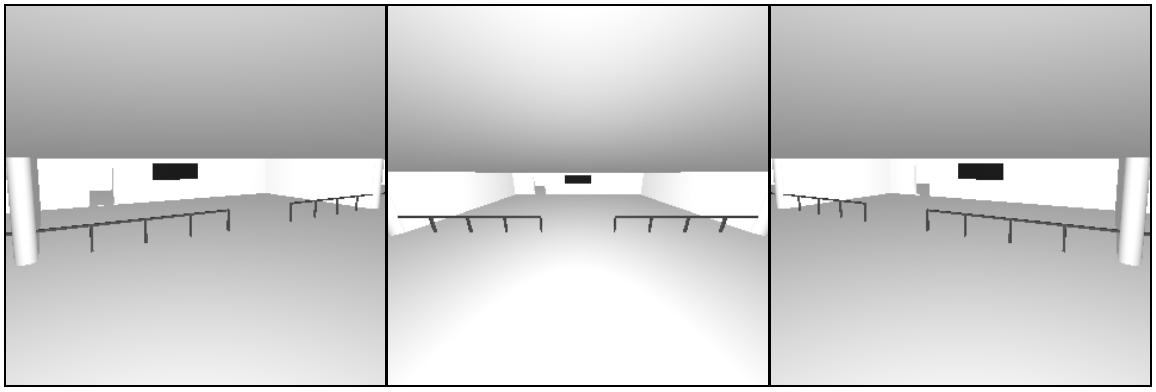


Figure 33. Three Camera View of Terminal without Desk

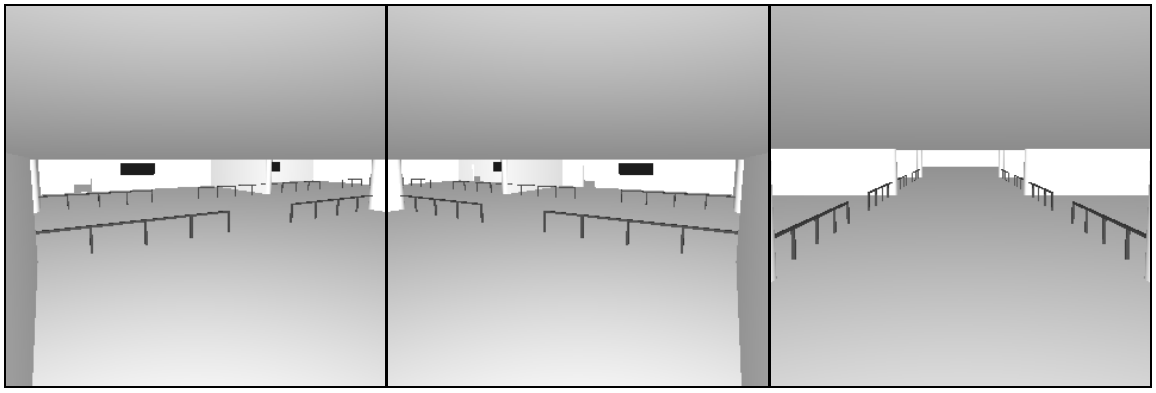


Figure 34. Three Camera View of Corridor

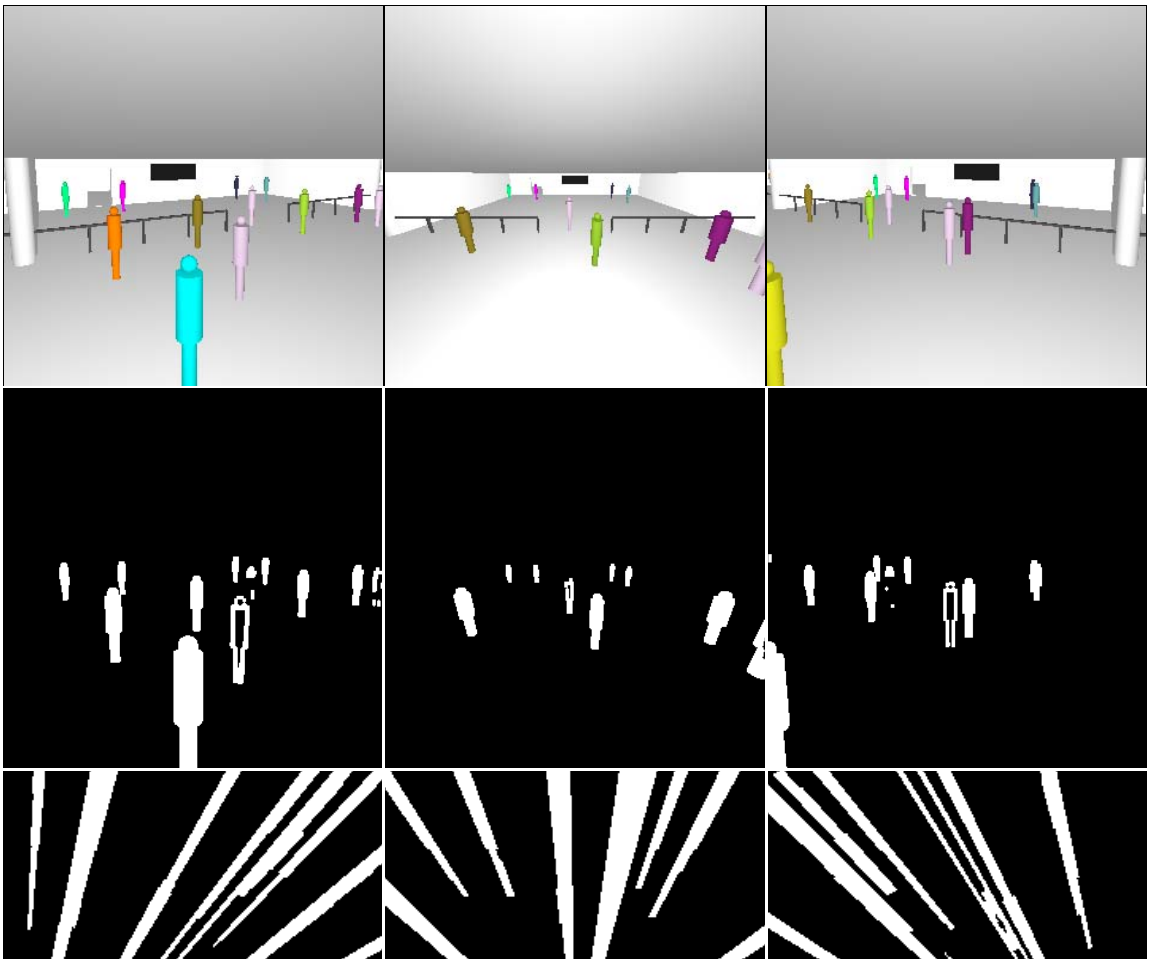


Figure 35. Processing of Three Camera Images for Terminal without Desk

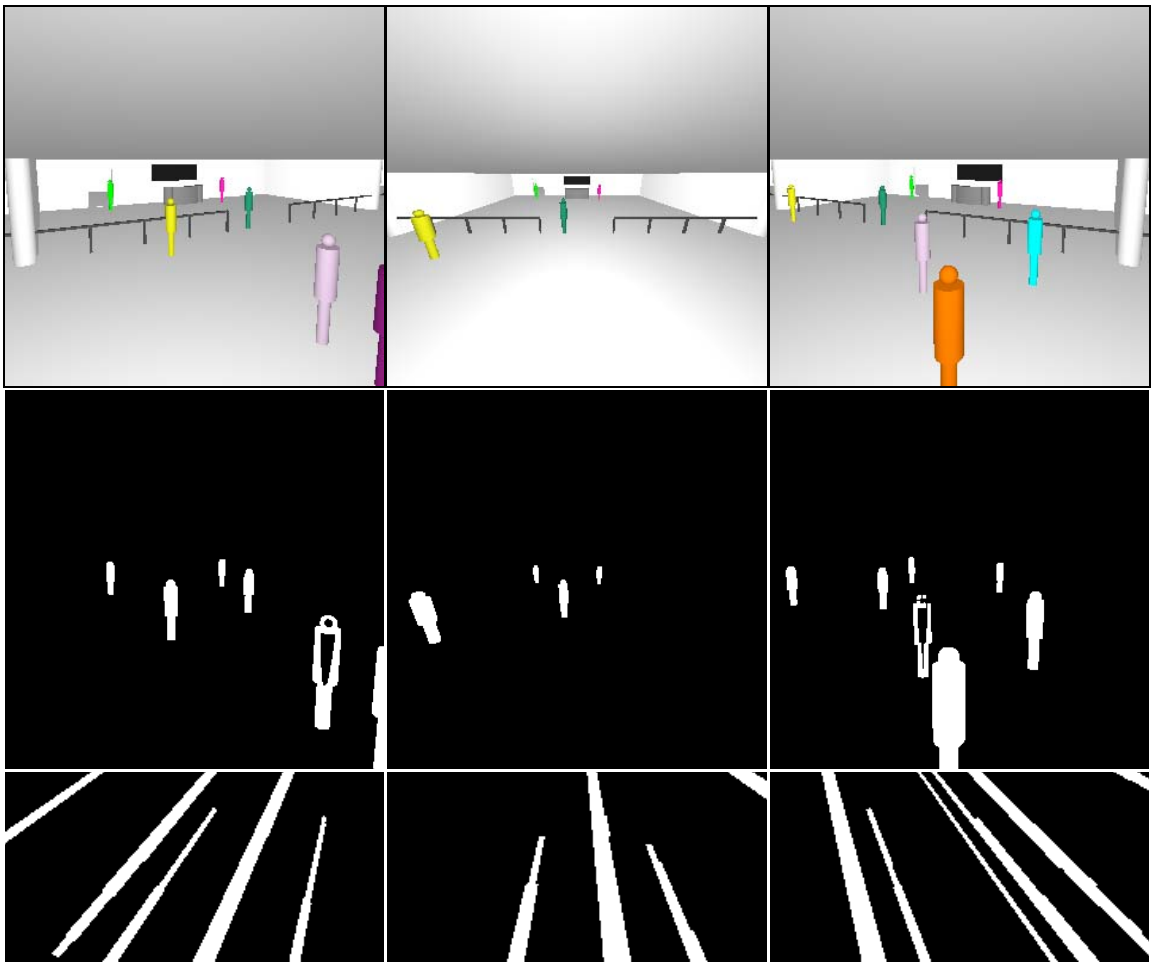


Figure 36. Processing of Three Camera Images for Terminal with Desk

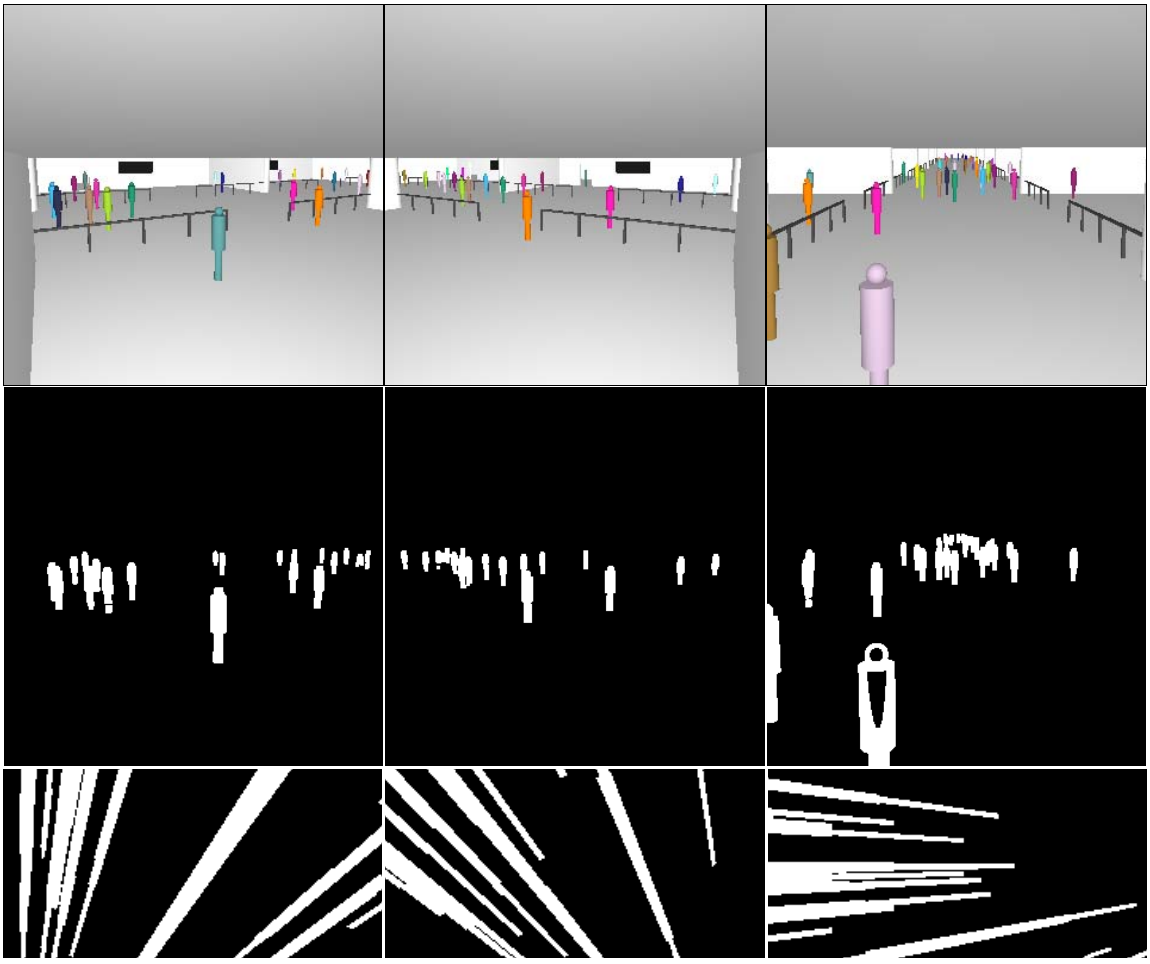


Figure 37. Processing of Three Camera Images for Corridor

E. Fusing LoS Sensor and Cameras

Fusion of the LoS and camera sensors is accomplished by a pixel-level AND operation of the rectified camera and LoS data. In a deployed system, the LoS sensor data is rectified to the camera so that the camera does not need to process every pixel. For this study, camera data is rectified to the LoS sensor plane so location data produced by the different sensor combinations can be compared. The AND operation is used because information from one sensor may invalidate detection hypotheses of another sensor. The cameras and LoS sensor Figure 38 on page 68 shows all sixteen combinations of three cameras and LoS sensor fusion for a terminal without a desk. Figure 39 on page 69 shows all sixteen combinations of three cameras and LoS sensor fusion for a terminal with a desk. Figure 40 on page 70 shows all sixteen combinations of three cameras and LoS sensor fusion for a portion of the corridor.

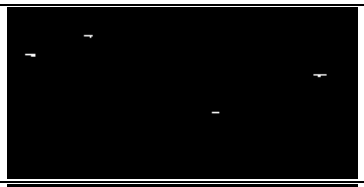
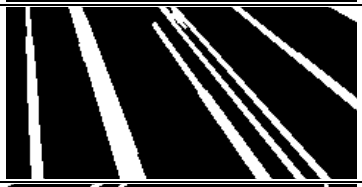
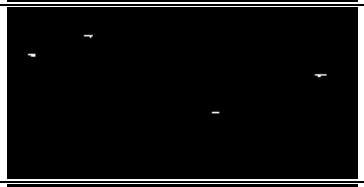
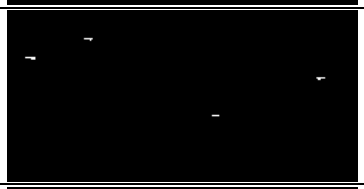
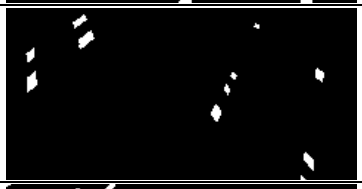
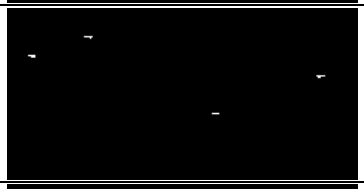
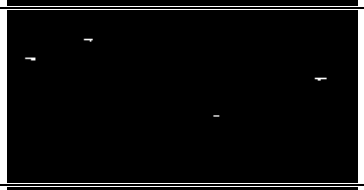
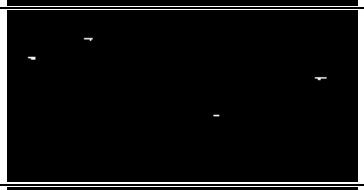
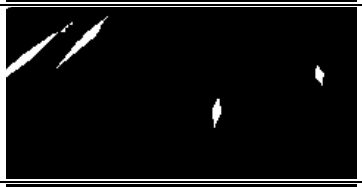
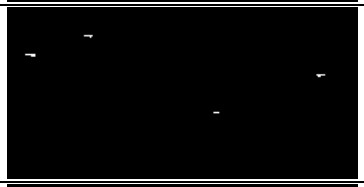
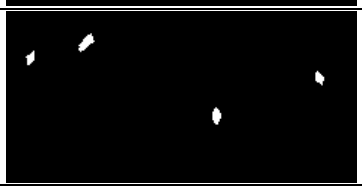
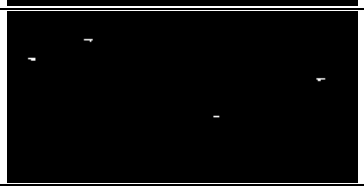
	Without LoS	With LoS
No Cameras		
Camera 1		
Camera 2		
Cameras 1 and 2		
Camera 3		
Cameras 1 and 3		
Cameras 2 and 3		
Cameras 1, 2 and 3		

Figure 38. Camera Image Data Fusion with LoS Sensor Data for Terminal With Desk

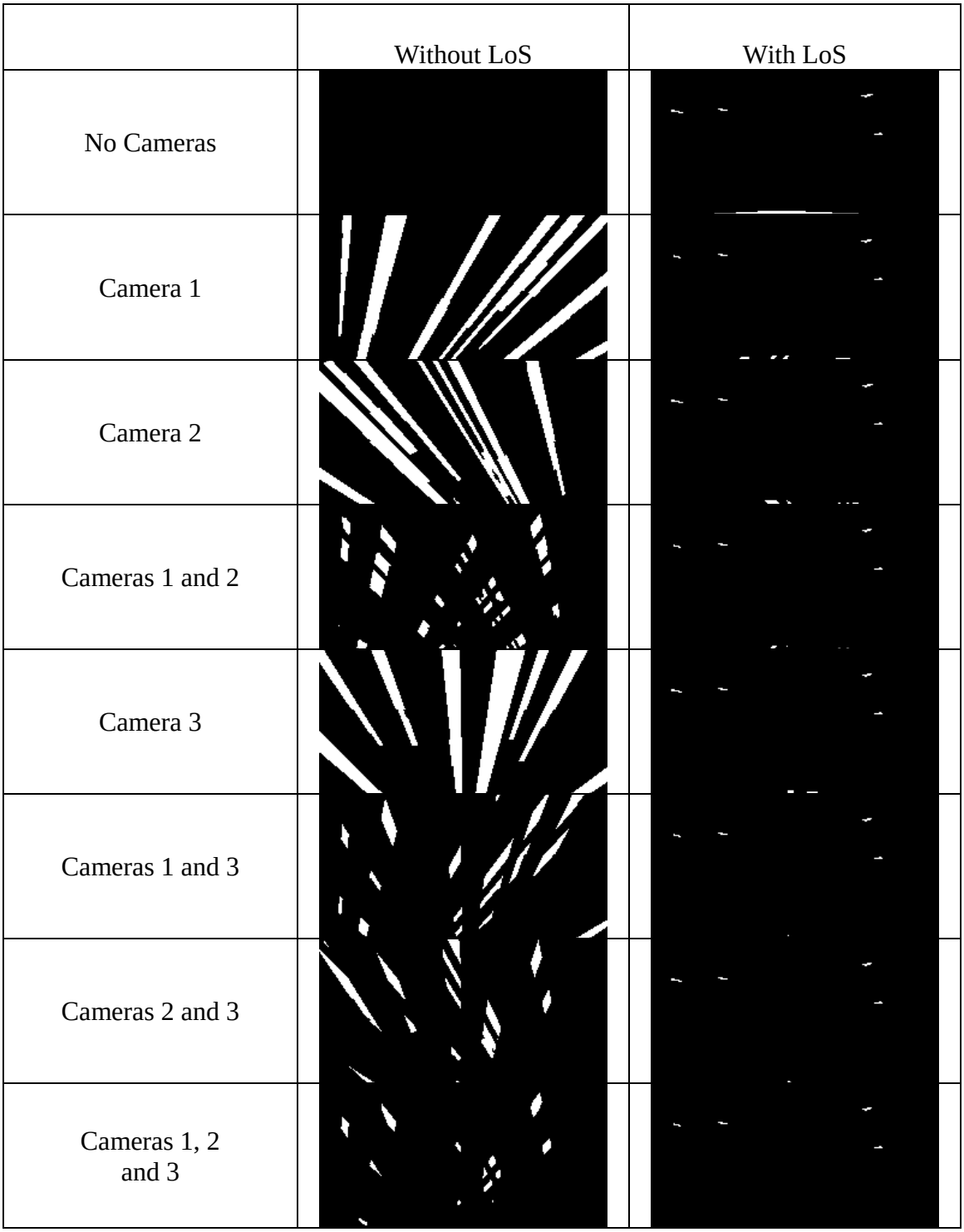


Figure 39. Camera Image Data Fusion with LoS Sensor Data for Terminal Without Desk

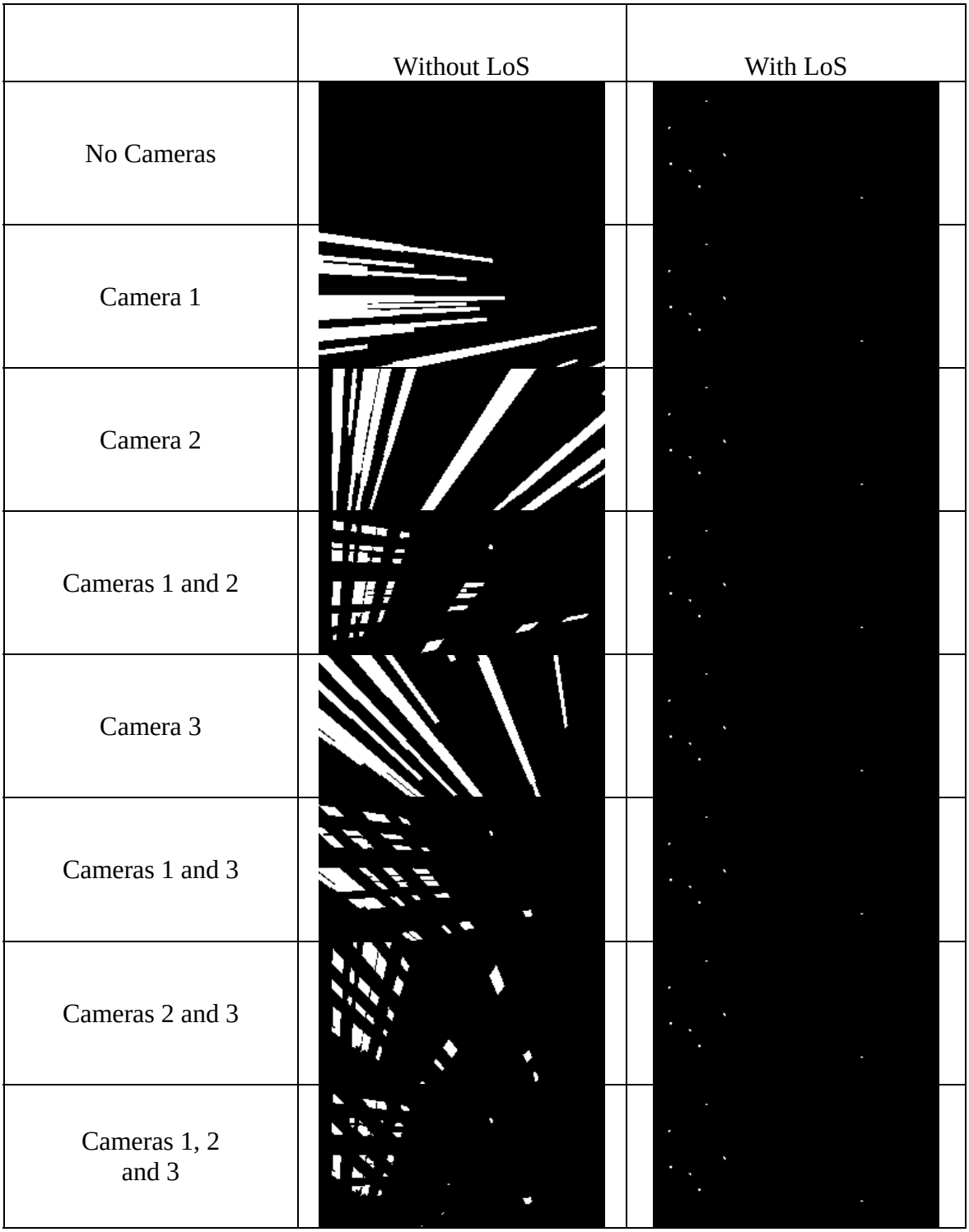


Figure 40. Camera Image Data Fusion with LoS Sensor Data for Corridor

F. Results from Simulation

Each simulation is repeated 50 times with each of 48 factor combinations for a total of 2,400 simulations. In addition, each simulation consists of three types of ROIs: a terminal with desk, a terminal without desk, and a portion of the corridor. Three performance metrics are collected and include the Hit Rate, False Alarm Rate, and precision.

G. LoS Sensor Characteristics

To characterize the hit rate of the LoS sensor with respect to the various factors Analysis of Variance (ANOVA) is used. A similar analysis is conducted when all three cameras are used without the LoS sensor. Finally, the results from the LoS study and the three cameras study are compared.

The two primary areas the LoS sensor are used in are the corridor and the terminals. The corridor is monitored by two LoS sensors cooperating with each other and LoS sensors in adjacent ROIs. In the terminals, each LoS sensor works independently.

Because of the differences in the deployment of the LoS sensor in the corridors and terminal, the two situations are examined together and separately with the presence or absence of the desk as a factor in the terminals.

1) LoS Sensor Hit Rate

Table 6 shows ANOVA of the hit rate for simulation results in which the LoS sensor is included. It can be seen from p-factor the number of people in the simulation, either considered alone or combined with other factors, is not statistically significant with respect to error. Table 6 also shows that the type of area contributes more to the variation than statistical error. Table 7 is an ANOVA of the hit rate for the LoS sensor that does

not include the number of people in the simulation as a factor so that an equation to estimate the mean HR can be derived of only significant factors.

Table 6. ANOVA of LoS Sensor Hit Rate

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Type Area	34.5	50.03%	2	17.3	2785	0
Number Cameras	0.732	1.06%	3	0.244	39.4	0
Number People	0.00917	0.01%	2	0.00458	0.739	0.477
Type Area*Number Cameras	0.481	0.70%	6	0.0802	12.9	1.60E-14
Type Area*Number People	0.0408	0.06%	4	0.0102	1.64	0.160
Number Cameras*Number People	0.00170	0.00%	6	0.000284	0.0457	0.99961
Type Area*Number Cameras*Number People	0.00723	0.01%	12	0.000602	0.0971	0.99997
Error	22.0	31.85%	3546	0.00620		
Total	69.0		3581			

Table 7. ANOVA of LoS Sensor Hit Rate Without Number of People Factor

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Type Area	34.5	50.04%	2	17.3	2795	0
Number Cameras	0.732	1.06%	3	0.244	39.5	0
Type Area*Number Cameras	0.481	0.70%	6	0.0801	13.0	1.48E-14
Error	22.1	31.96%	3570	0.00618		
Total	69.0		3581			

Table 7 shows the mean hit rate of the LoS sensor is determined by both the type of area in which the sensor is used and the number of cameras it is used with.

a) LoS Hit Rate for Corridor

ANOVA of the hit rate for the corridor deployed LoS sensor given a number of cameras is shown in Table 8, with the resulting coefficients shown in Table 9.

Table 8. ANOVA of Corridor LoS Sensor Hit Rate

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Number Cameras	1.11	15.33%	3	0.369	71.8	0
Error	6.11	84.67%	1190	0.00514		
Total	7.22		1193			

Table 9. Number of Camera Coefficients for Corridor LoS Sensor Mean Hit Rate

Number Cameras	Coefficient
0	0.0634
1	0.00286
2	-0.0254
3	-0.0409

The constant coefficient of the mean hit rate is 0.901. The equation to estimate the mean hit rate for the corridor LoS sensor is then:

$$\text{Corridor LoS HR} = 0.901 + \textit{Number Cameras Coefficient} \quad (1)$$

Table 10 and Figure 41 contain the mean and 95% confidence interval for the hit rate of the corridor LoS sensor given the number of cameras and Figure 42 shows a box plot of the distribution. It can be seen that the confidence intervals of the levels for two and three cameras overlap, and are therefore not statistically different at the 95% confidence level. However, the levels for zero and one are statistically different from each other and from the levels for two or three cameras.

Table 10. Confidence Interval of Corridor LoS Sensor Mean Hit Rate Given Number of Cameras

Cameras	Mean	Error (95%)	Confidence Interval
0	0.965	0.00361	[0.968, 0.961]
1	0.904	0.00327	[0.907, 0.901]
2	0.876	0.00366	[0.880, 0.872]
3	0.860	0.00678	[0.867, 0.854]

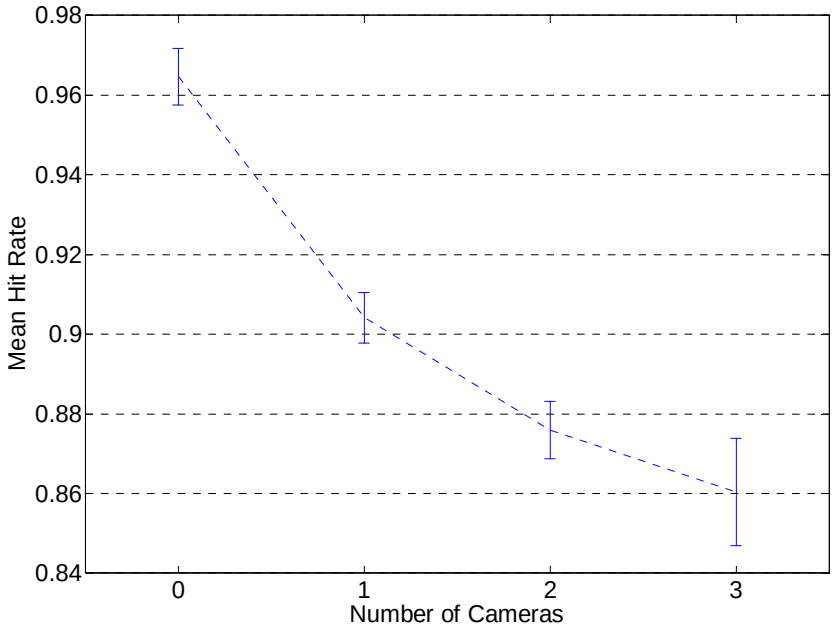


Figure 41. Confidence Interval of Corridor LoS Sensor Hit Rate

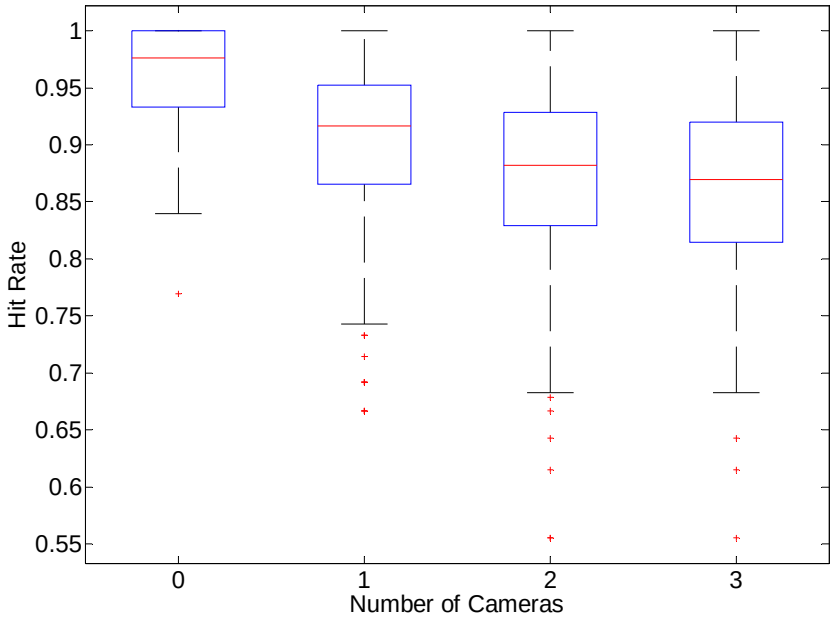


Figure 42. Distribution of Corridor LoS Sensor Hit Rate

b) LoS Hit Rate for Terminal

ANOVA of the hit rate for the terminal deployed LoS sensor is shown in Table 11. It can be seen by the p-factor the number of cameras and the presence of the desk are statistically significant in determining the mean hit rate of the LoS sensor, but their combined effects are not statistically significant.

Table 11. ANOVA of Terminal LoS Sensor Hit Rate

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Number Cameras	0.0990	0.16%	3	0.0330	4.93	0.00206
Desk	33.3	55.00%	1	33.3	4965	0
Number Cameras*Desk	0.00737	0.01%	3	0.00246	0.367	0.777
Error	16.0	26.36%	2380	0.00670		
Total	60.5		2387			

Table 12 is the one-way ANOVA of the hit rate for the terminal LoS sensor, and Table 13 and Table 14 are the resulting coefficients. The constant coefficient of the mean hit rate is 0.855.

Table 12. One-Way ANOVA of Terminal LoS Sensor Hit Rate

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Number Cameras	0.0990	0.16%	3	0.0330	4.93	0.00205
Desk	44.4	73.46%	1	44.4	6637	0
Error	16.0	26.37%	2383	0.00670		
Total	60.5		2387			

Table 13. Number of Camera Coefficients for Terminal LoS Sensor Mean Hit Rate

Number Cameras	Coefficient
0	0.0114
1	0.00350
2	-0.00390
3	-0.0110

Table 14. Desk Coefficients for Terminal LoS Sensor Mean Hit Rate

Presence of Desk	Coefficient
With Desk	-0.136
Without Desk	0.136

The equation to estimate the mean hit rate for the terminal LoS sensor is then:

$$\text{Terminal LoS HR} = 0.855 + \textit{Number Cameras Coefficient} + \textit{Presence of Desk Coefficient} \quad (2)$$

Table 15 and Figure 43 contain the mean and 95% confidence interval of the hit rate for the terminal LoS sensor given the number of cameras and Figure 44 shows a box plot of the distribution. It can be seen that the confidence intervals of the levels for zero, one, two, and three cameras all overlap, and therefore are not statistically different at the 95% confidence level.

Table 15. Confidence Intervals of Terminal LoS Sensor Mean Hit Rate Given Number of Cameras

Cameras	Mean	Error (95%)	Confidence Interval
0	0.867	0.0175	[0.842, 0.849]
1	0.859	0.0104	[0.869, 0.848]
2	0.851	0.0105	[0.862, 0.841]
3	0.844	0.0186	[0.863, 0.826]

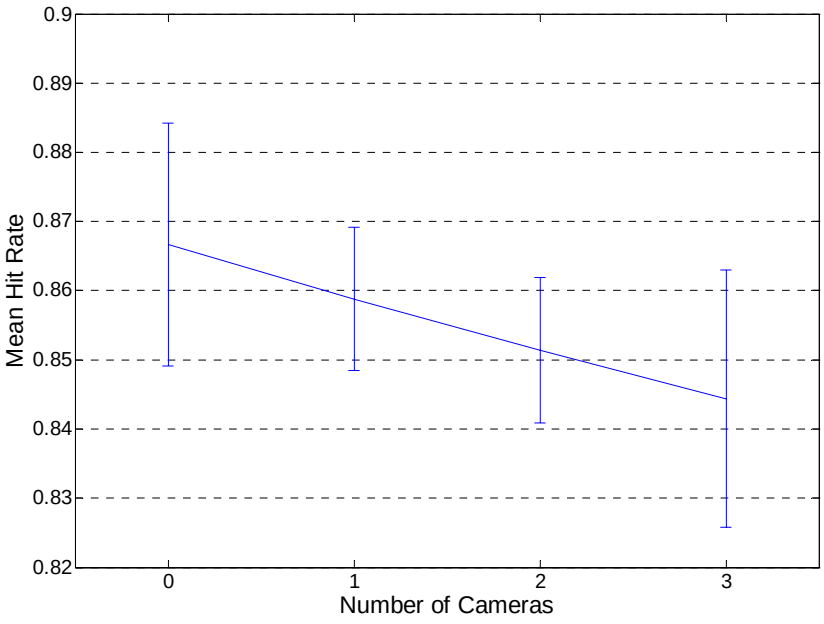


Figure 43. Confidence Interval of Terminal LoS Sensor Hit Rate Given Number of Cameras

Table 16 and Figure 45 contain the mean and 95% confidence interval of the terminal LoS sensor given the presence or absence of the desk and Figure 46 shows a box plot of the distribution. The confidence intervals of the levels do not overlap. The hit rate of the LoS sensor in a terminal with a desk is statistically different than the hit rate of the LoS sensor in a terminal without a desk.

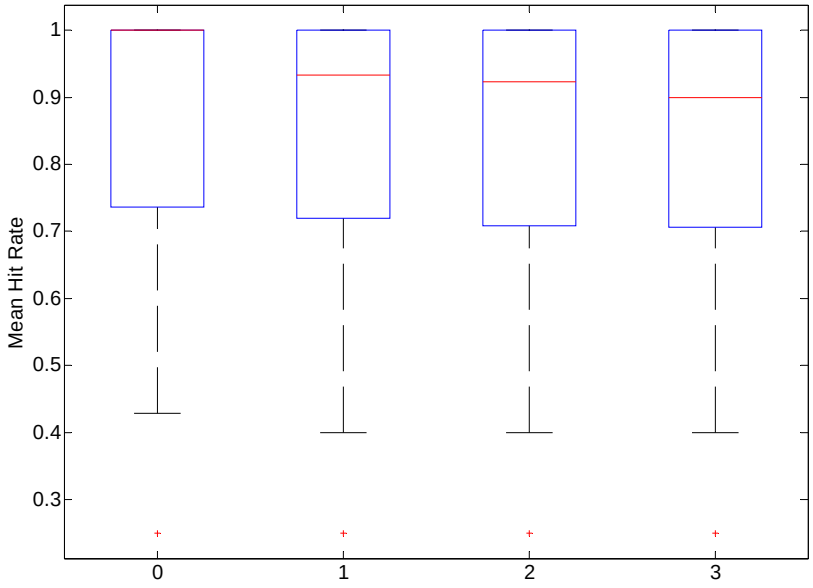


Figure 44. Distribution of Terminal LoS Sensor Hit Rate Given Number of Cameras

Table 16. Confidence Intervals of Terminal LoS Sensor Mean Hit Rate given Presence of Desk

	Mean	Error (95%)	Confidence Interval
Without Desk	0.719	0.00640	[0.725, 0.712]
With Desk	0.992	0.00155	[0.993, 0.990]

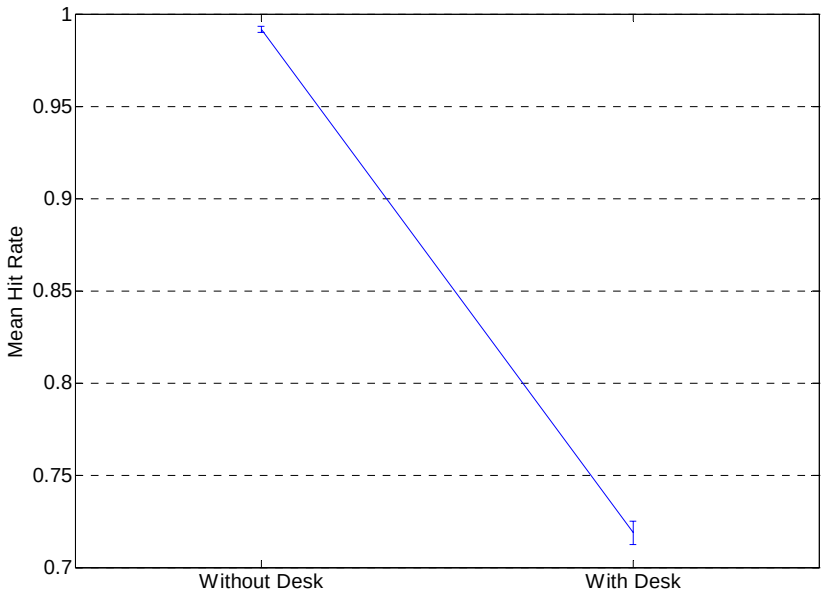


Figure 45. Confidence Interval of Terminal LoS Sensor Hit Rate Given Presence of Desk

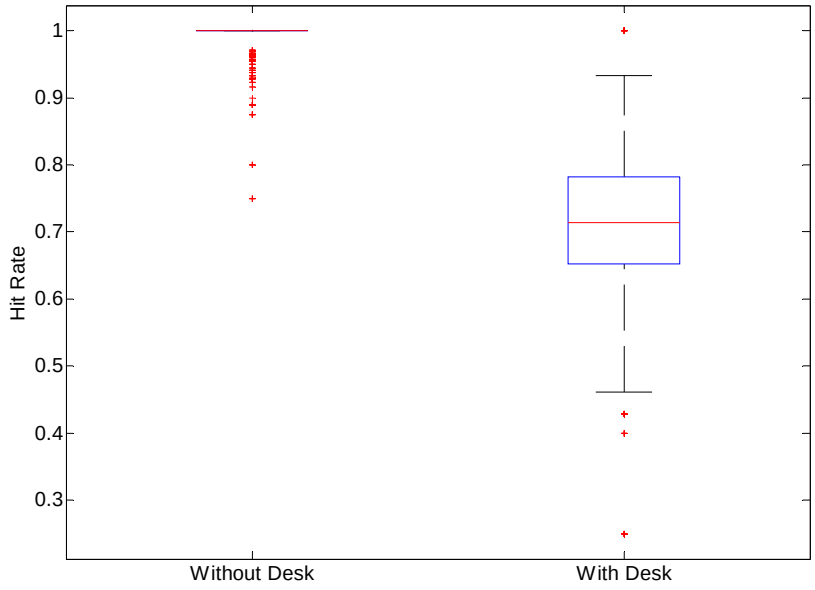


Figure 46. Distribution of Terminal LoS Sensor Hit Rate Given Presence of Desk

2) LoS False Alarm Rate

Table 17 shows the results of ANOVA of the false alarm rate for simulation results in which the LoS sensor is included. It can be seen by the p-factor from the table that the effects of the area, number of cameras, and number of people are statistically significant independently and in combination.

Table 17. ANOVA of LoS Sensor False Alarm Rate

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Type Area	1186	14.71%	2	593	392	0
Number Cameras	193	2.40%	3	64.4	42.5	0
Number People	327	4.06%	2	163	108	0
Type Area*Number Cameras	366	4.54%	6	61.0	40.3	0
Type Area*Number People	277	3.44%	4	69.3	45.7	0
Number Cameras*Number People	32.2	0.40%	6	5.37	3.55	0.00167
Type Area*Number Cameras *Number People	49.6	0.62%	12	4.14	2.73	0.00109
Error	5370	66.63%	3546	1.51		
Total	8059		3581			

a) LoS False Alarm Rate for Corridor

The ANOVA of the false alarm rate for the corridor deployed LoS sensor given a number of cameras is shown in Table 18, with the resulting coefficients shown in Table 19, Table 20, and Table 21. The constant coefficient of the mean hit rate is 1.64.

Table 18. ANOVA of Corridor LoS Sensor False Alarm Rate

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Number Cameras	549	8.99%	3	183	45.4	0
Number People	573	9.38%	2	287	71.0	0
Number Cameras*Number People	75.6	1.24%	6	12.6	3.13	0.00482
Error	4768	78.06%	1182	4.03		
Total	6108		1193			

Table 19. Number of Camera Coefficients of Corridor LoS Sensor Mean False Alarm Rate

Number of Cameras	Coefficient
0	1.42
1	0.0504
2	-0.574
3	-0.894

Table 20. Number of People Coefficients of Corridor LoS Sensor Mean False Alarm Rate

Number People	Coefficient
30	-0.942
60	-0.0723
90	1.01

Table 21. Number of Camera Combined With Number of People Coefficients of Corridor LoS Sensor Mean False Alarm Rate

Number People	Number Cameras	Coefficient
30	0	-0.618
	1	0.0523
	2	0.565
	3	-0.0637
60	0	-0.0737
	1	0.137
	2	0.228
	3	0.0178
90	0	-0.246
	1	0.454
	2	0.00358
	3	-0.457

The equation to estimate the mean false alarm rate for the corridor LoS sensor is:

$$\text{Corridor LoS FAR} = 1.64 + \textit{Number Cameras Coefficient} + \textit{Number People Coefficient} + \textit{Number Cameras Combined With Number People Coefficient.} \quad (3)$$

Table 22 and Figure 47 contain the mean and 95% confidence interval of the terminal LoS sensor given the presence or absence of the desk and Figure 48 shows a box plot of the distribution.

Table 22. Confidence Intervals of Terminal LoS Sensor Mean False Alarm Rate

Number People	Number Cameras	Mean	Error (95%)	Confidence Interval
30	0	1.50	0.829	[2.33, 0.671]
	1	0.687	0.278	[0.965, 0.409]
	2	0.353	0.115	[0.469, 0.238]
	3	0.260	0.160	[0.420, 0.100]
60	0	3.04	0.936	[3.98, 2.13]
	1	1.55	0.296	[1.84, 1.25]
	2	1.01	0.189	[1.20, 0.825]
	3	0.680	0.272	[0.952, 0.408]
90	0	4.64	1.168	[5.81, 3.47]
	1	2.84	0.448	[3.29, 2.40]
	2	1.84	0.303	[2.14, 1.53]
	3	1.31	0.372	[1.68, 0.934]

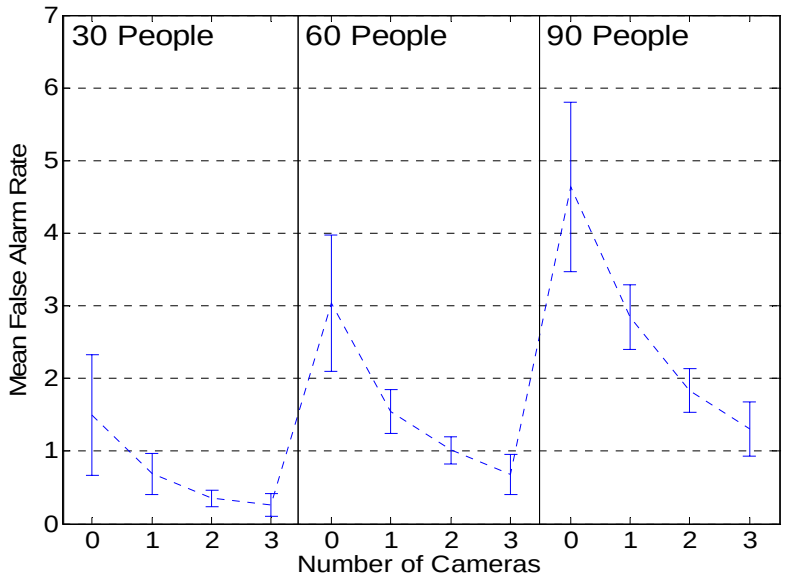


Figure 47. Confidence Interval of Corridor LoS Sensor False Alarm Rate Given Number of Cameras and Number of People

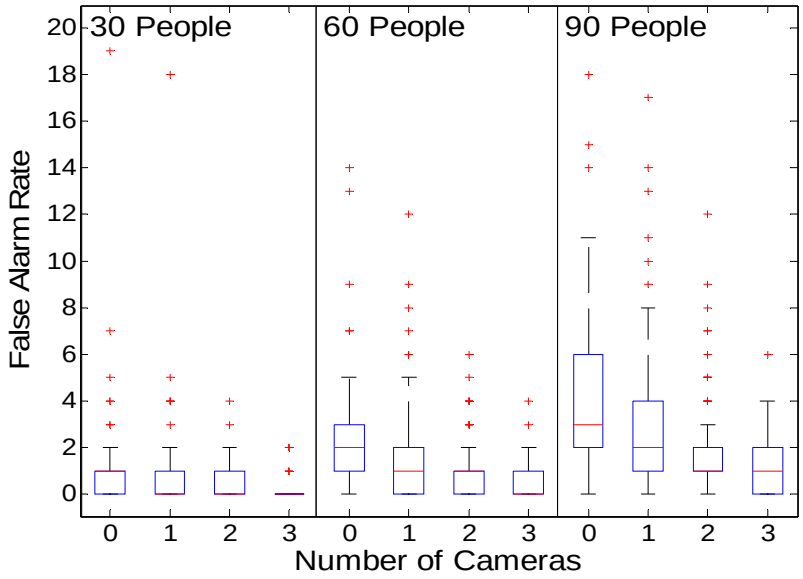


Figure 48. Distribution of Corridor LoS Sensor False Alarm Rate Given Number of Cameras and Number of People

Figure 48 shows that for 30, 60 and 90 people, there is not a statistically significant difference at the 95% confidence level in the mean false alarm rate between the two and three camera levels. For 30 people, there is not a statistically significant difference between zero and one cameras, or between one and two cameras. However, there is a statistically significant difference between zero and two or more cameras and between one and three cameras. For 60 and 90 people, there is a statistically significant difference between zero, one and two cameras. There is also a statistically significant difference between zero, one and three cameras.

Figure 48 also shows that for zero cameras, there is no statistically significant difference in the mean false alarm rate between 30 and 60 people, and between 60 and 90 people. However, there is a statistically significant difference between 30 and 90 people for zero cameras. For one and two cameras, there is a statistically significant difference between 30, 60 and 90 people. For three cameras, there is a statistically significant difference between 30 and 90 people, but not for between 30 and 60 people, or between 60 or 90 people.

b) LoS False Alarm Rate for Terminal

ANOVA of the false alarm rate for the terminal deployed LoS sensor given a number of cameras is shown in Table 23. It can be seen by the p-factor that the number of cameras is not statistically significant when considered alone or combined with the number of cameras. For Table 24, the number of cameras and number of cameras combined with the number of people effects are removed. Table 25, Table 26, Table 27, Table 28, and Table 29 provide the coefficients for estimating the mean false alarm rate

for the terminal LoS sensor. The constant coefficient for estimating the false alarm rate is 0.247.

Table 23. ANOVA of Terminal LoS Sensor False Alarm Rate

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Desk	21.5	3.16%	1	21.5	84.3	0
Number Cameras	0.655	0.10%	3	0.218	0.857	0.463
Number People	27.5	4.05%	2	13.7	54.0	0
Desk*Number Cameras	9.72	1.43%	3	3.24	12.7	2.99E-08
Desk*Number People	3.33	0.49%	2	1.66	6.53	0.00148
Number Cameras*Number People	1.87	0.28%	6	0.312	1.23	0.289
Desk*Number Cameras*Number People	4.35	0.64%	6	0.725	2.84	0.00920
Error	602	88.71%	2364	0.255		
Total	679		2387			

Table 24. ANOVA of Terminal LoS Sensor False Alarm Rate Without Number of Cameras

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Desk	21.6	3.18%	1	21.6	83.7	0
Number People	27.5	4.06%	2	13.8	53.4	0
Number Cameras*Desk	4.51	0.67%	3	1.50	5.84	0.000572
Desk*Number People	3.37	0.50%	2	1.69	6.53	0.00148
Number Cameras*Desk*Number People	2.20	0.32%	6	0.366	1.42	0.203
Error	612	90.17%	2373	0.258		
Total	679		2387			

Table 25. Presence of Desk Coefficients of Terminal LoS Sensor Mean False Alarm Rate

Desk	Coefficient
With	0.110
Without	-0.110

Table 26. Number of People Coefficients of Terminal LoS Sensor Mean False Alarm Rate

Number People	Coefficient
30	-0.133
60	-0.0324
90	0.166

Table 27. Number of Cameras with Presence of Desk Coefficients of Terminal LoS Sensor Mean False Alarm Rate

Number Cameras	Desk	Coefficient
0	With	0.0662
	Without	-0.0662
1	With	-0.0320
	Without	0.0320
2	With	-0.0600
	Without	0.0600
3	With	0.0258
	Without	-0.0258

Table 28. Number of People with Presence of Desk Coefficients of Terminal LoS Sensor Mean False Alarm Rate

Number People	Desk	Coefficient
30	With	-0.0597
	Without	0.0178
60	With	0.0419
	Without	0.0597
90	With	-0.0178
	Without	-0.0419

Table 29. Number of People, Number of Cameras and Presence of Desk Coefficients of Terminal LoS Sensor Mean False Alarm Rate

Number People	Number Cameras	Desk	Coefficient
30	0	With	-0.0170
		Without	-0.0316
	1	With	0.0487
		Without	0.0170
	2	With	0.0316
		Without	-0.0487
	3	With	0.00785
		Without	0.00993
60	0	With	-0.0178
		Without	-0.00785
	1	With	-0.00993
		Without	0.0178
	2	With	0.0341
		Without	0.0371
	3	With	-0.0712
		Without	-0.0341
90	0	With	-0.0371
		Without	0.0712
	1	With	-0.0249
		Without	-0.0154
	2	With	0.0403
		Without	0.0249
	3	With	0.0154
		Without	-0.0403

The equation to estimate the mean false alarm rate for the terminal LoS sensor is:

$$\text{Terminal LoS FAR} = 0.247 + \textit{Presence of Desk Coefficient} + \textit{Number People Coefficient} + \textit{Number Cameras with Presence of Desk Coefficient} + \textit{Number People with Presence of Desk Coefficient} \quad (4)$$

+ Number People, Number Cameras, and Presence of Desk Coefficient

Table 30 and Figure 49 contain the mean and 95% confidence interval of the terminal LoS sensor given the presence or absence of the desk and Figure 50 shows a box plot of the distribution.

Table 30. Confidence Intervals of Terminal LoS Sensor Mean Hit Rate Given Presence of Desk

Number People	Desk	Number Cameras	Mean	Error (95%)	Confidence Interval
30	With	0	0.260	0.138	[0.398, 0.122]
		1	0.120	0.0618	[0.182, 0.0582]
		2	0.133	0.0610	[0.194, 0.0723]
		3	0.140	0.0996	[0.240, 0.0404]
	Without	0	0.00	0.00	[0.00, 0.00]
		1	0.0667	0.0404	[0.107, 0.0263]
		2	0.0867	0.0527	[0.139, 0.340]
		3	0.100	0.0861	[0.186, 0.0139]
60	With	0	0.400	0.172	[0.572, 0.228]
		1	0.327	0.0998	[0.426, 0.227]
		2	0.300	0.0892	[0.389, 0.211]
		3	0.340	0.158	[0.498, 0.182]
	Without	0	0.00	0.00	[0.00, 0.00]
		1	0.100	0.0520	[0.152, 0.0480]
		2	0.107	0.0565	[0.163, 0.0501]
		3	0.140	0.0996	[0.240, 0.0404]
90	With	0	0.860	0.215	[1.07, 0.645]
		1	0.561	0.137	[0.698, 0.424]
		2	0.401	0.0985	[0.500, 0.303]
		3	0.429	0.185	[0.614, 0.243]
	Without	0	0.00	0.00	[0.00, 0.00]
		1	0.399	0.110	[0.509, 0.289]
		2	0.320	0.0934	[0.413, 0.226]
		3	0.327	0.159	[0.486, 0.167]

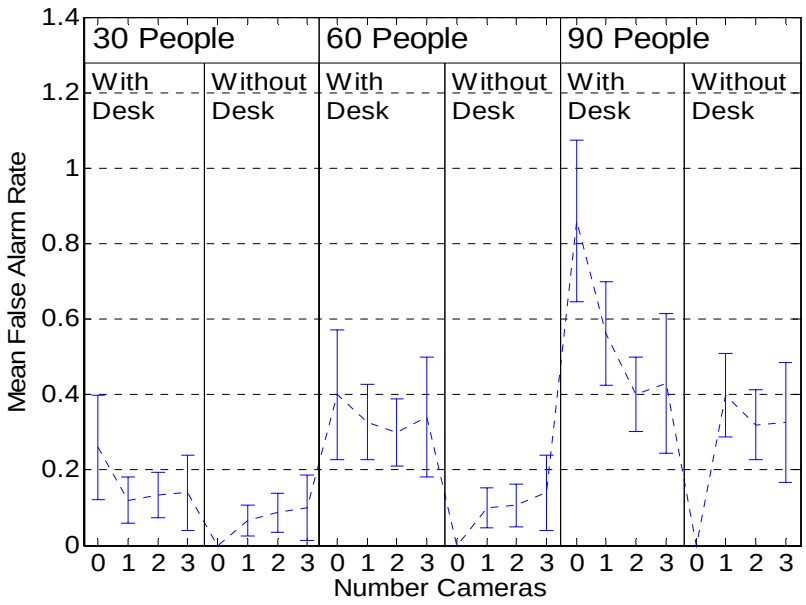


Figure 49. Confidence Interval of Terminal Deployed LoS Sensor False Alarm Rate Given Number of Cameras, Number of People, and Presence of Desk

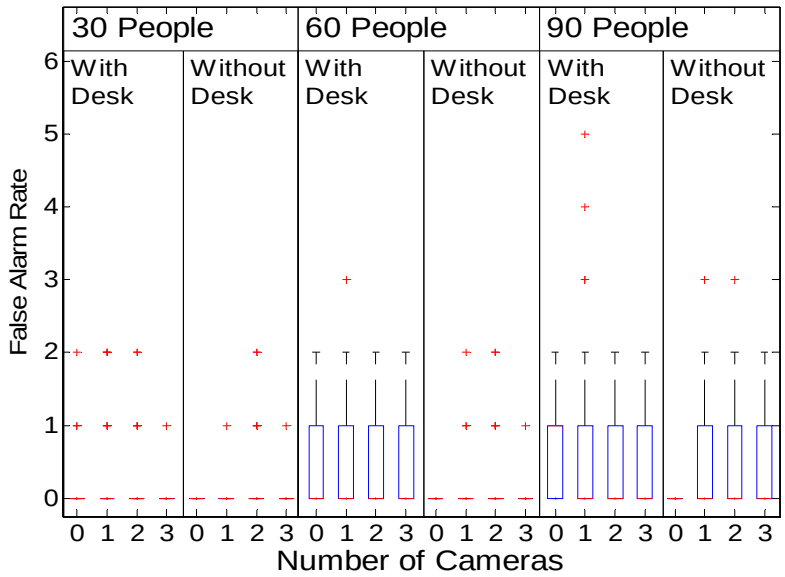


Figure 50. Distribution of Terminal Deployed LoS Sensor False Alarm Rate Given Number of Cameras, Number of People, and Presence of Desk

Figure 50 shows that for 30, 60 and 90 people, no false alarms occurred when the LoS sensor is used alone in a terminal without a desk. However, in the terminal with a desk, the false alarm rate for the LoS sensor used alone is statistically greater for 90 people than for 30 people. The LoS sensor used with one camera in a terminal with a desk is also statistically greater than with 90 people than with 30 people.

3) LoS Sensor Precision

ANOVA of the precision, measured in square feet, for the LoS sensor given a number of cameras is shown in Table 31. The p-factor indicates the three-level interaction between the type of area, number of cameras, and number of people is not a statistically significant with respect to error. On page 92, Table 32 is a two-way ANOVA of the precision for the LoS sensor.

Table 31. ANOVA of LoS Sensor Precision

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Type Area	55951	51.87%	2	27976	4120	0
Number Cameras	10509	9.74%	3	3503	516	0
Number People	552	0.51%	2	276	40.6	0
Type Area*Number Cameras	798	0.74%	6	133	19.6	0
Type Area*Number People	114	0.11%	4	28.6	4.21	0.00210
Number Cameras*Number People	110	0.10%	6	18.4	2.71	0.0126
Type Area*Number Cameras*Number People	84.7	0.08%	12	7.06	1.04	0.409
Error	24077	22.32%	3546	6.79		
Total	107860		3581			

Table 32. Two-Way ANOVA of LoS Sensor Precision

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Type Area	55961	51.88%	2	27980	4120	0
Number Cameras	10509	9.74%	3	3503	516	0
Number People	552	0.51%	2	276	40.6	0
Type Area*Number Cameras	797	0.74%	6	133	19.6	0
Type Area*Number People	112	0.10%	4	28.1	4.14	0.00239
Number Cameras*Number People	110	0.10%	6	18.4	2.71	0.0126
Error	24162	22.40%	3558	6.79		
Total	107860		3581			

a) LoS Sensor Precision for Corridor

The ANOVA of the precision for the corridor deployed LoS sensor given a number of cameras is shown in Table 33. As can be seen by the p-factor, the two-way interaction of the effects of the number of cameras and the number of people is not statistically significant. Therefore, the effects due to the number of cameras and the number of people are considered separately.

Table 33. ANOVA of Corridor LoS Sensor Precision

Source	**Sum Sq.**	**% Var**	**d.f.**	**Mean Sq.**	**F**	**Prob>F**
Number Cameras	5.24	10.40%	3	1.75	46.6	0
Number People	0.555	1.10%	2	0.278	7.41	0.000633
Number Cameras*Number People	0.0698	0.14%	6	0.0116	0.310	0.932
Error	44.3	87.97%	1182	0.0375		
Total	50.3		1193			

Table 34 is the one-way ANOVA of the precision for the corridor deployed LoS sensor for the number of people and the number of cameras. Table 35 and Table 36

provide the coefficients for estimating the mean precision for the corridor LoS sensor. The constant coefficient for estimating the precision is 1.19 square feet.

Table 34. One-Way ANOVA of Corridor LoS Sensor Precision

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Number Cameras	5.24	10.41%	3	1.75	46.8	0
Number People	0.739	1.47%	2	0.370	9.90	0.0000545
Error	44.4	88.11%	1188	0.0373		
Total	50.3		1193			

Table 35. Number of Cameras Coefficients of Corridor LoS Sensor Mean Precision

Number Cameras	Coefficient
0	0.138
1	0.00391
2	-0.0598
3	-0.0822

Table 36. Presence of Desk Coefficients of Corridor LoS Sensor Mean Precision

Number People	Coefficient
30	**-0.0351**
60	**0.0188**
90	**0.0163**

The equation to estimate the mean precision for the terminal LoS sensor is:

$$\text{Corridor LoS Precision} = 1.19 + \textit{Number People Coefficient} + \textit{Number Cameras Coefficient} \quad (5)$$

Table 37 and Figure 51 contain the mean and 95% confidence interval of the terminal LoS sensor given the presence or absence of the desk and Figure 52 shows a box plot of the distribution.

Table 37. Confidence Intervals of Corridor LoS Sensor Precision Given Number of Cameras

Number Cameras	Mean	Error (95%)	Confidence Interval
0	1.33	0.0527	[1.38, 1.27]
1	1.19	0.0197	[1.21, 1.17]
2	1.13	0.0118	[1.14, 1.12]
3	1.11	0.0176	[1.12, 1.09]

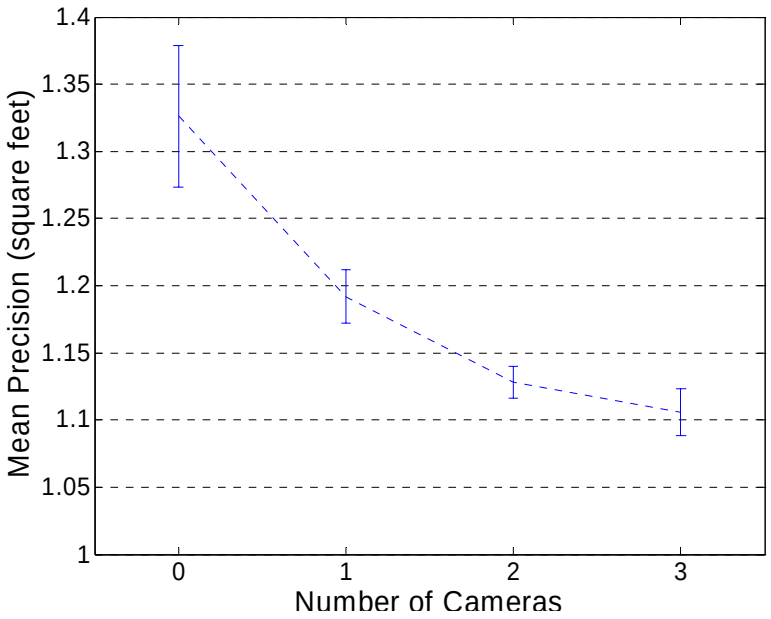


Figure 51. Confidence Interval of Corridor LoS Sensor Precision Given Number of Cameras

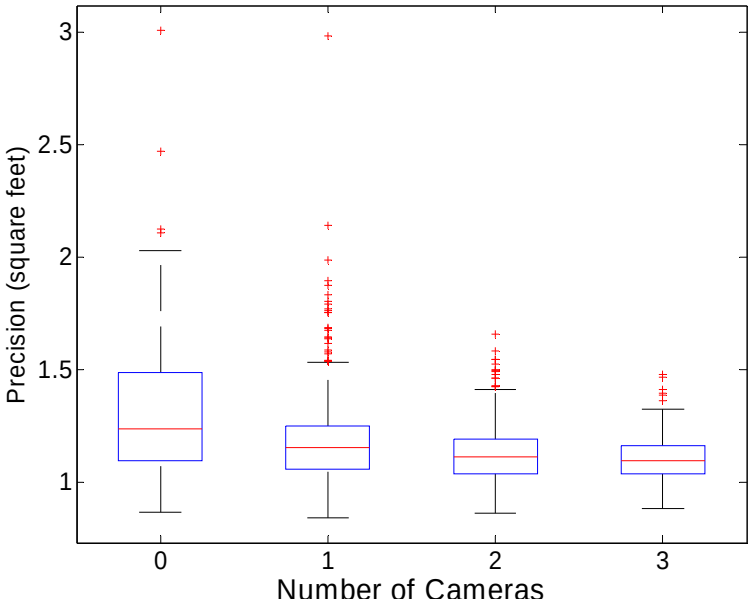


Figure 52. Distribution of Corridor LoS Sensor Precision Given Number of Cameras

From Figure 54, shows the confidence intervals of the levels for two and three cameras overlap, and are therefore not statistically different. However, levels for zero, one and two cameras do not overlap, and therefore are statistically different.

Table 38 and Figure 53 contain the mean and 95% confidence interval of the terminal LoS sensor given the presence or absence of the desk and Figure 54 shows a box plot of the distribution.

Table 38. Confidence Intervals of Corridor LoS Sensor Precision given Number of People

Number People	Mean	Error (95%)	Confidence Interval
30	1.14	0.0243	[1.16, 1.11]
60	1.19	0.0202	[1.21, 1.17]
90	1.19	0.0144	[1.21, 1.18]

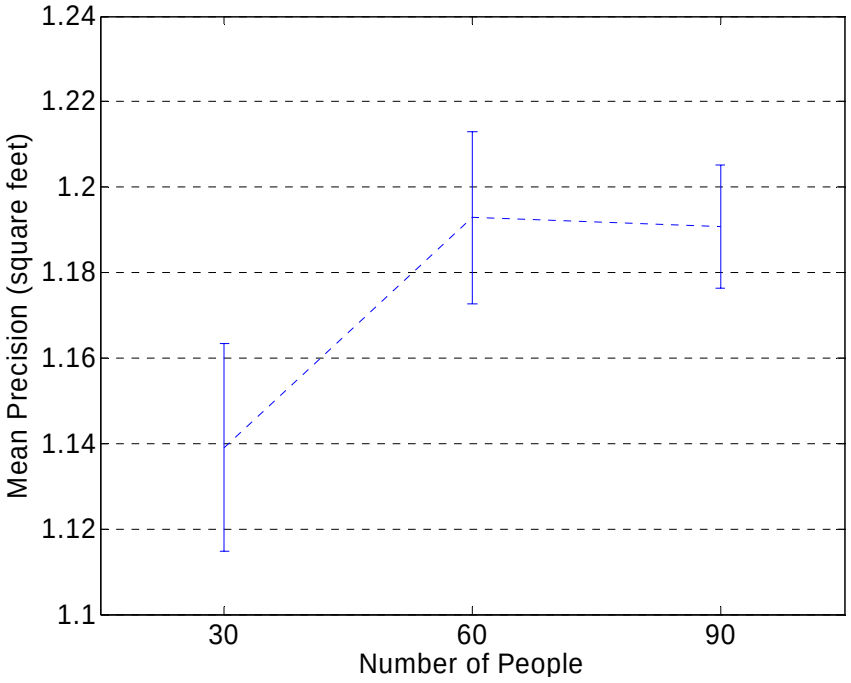


Figure 53. Confidence Interval of Corridor LoS Sensor Precision Given Number of People

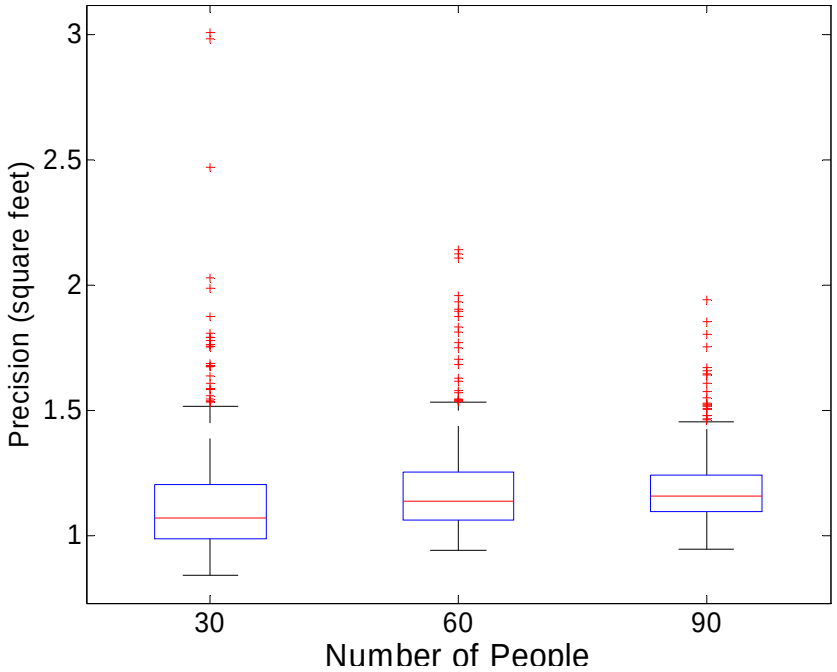


Figure 54. Distribution of Corridor Deployed LoS Sensor Precision Given Number of People

From Figure 54, shows the confidence intervals as the number of people increases from 30 to 60 do not overlap, and the differences in the accuracies are therefore statistically different. However, as the number of people is further increased from 60 to 90 people, the confidence levels do overlap, indicating that the precision between these two levels is not statistically different.

b) LoS Sensor Precision for Terminal

ANOVA of the precision for the terminal deployed LoS sensor given a number of cameras is shown in Table 39. As can be seen by the p-factor, the three-way interaction

of the effects, and the interaction of the effects from the presence of the desk and the number of cameras, is not statistically different.

Table 39. ANOVA of Terminal LoS Sensor Precision

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Desk	10.7	10.02%	1	10.7	507	0
Number Cameras	38.9	36.47%	3	13.0	615	0
Number People	1.72	1.61%	2	0.86	40.8	0
Desk*Number Cameras	0.0651	0.06%	3	0.0217	1.03	0.378
Desk*Number People	0.328	0.31%	2	0.164	7.78	0.000427
Number Cameras*Number People	0.572	0.54%	6	0.0953	4.53	0.000143
Desk*Number Cameras*Number People	0.120	0.11%	6	0.0200	0.951	0.457
Error	49.8	46.71%	2364	0.0211		
Total	107		2387			

Table 40 is an ANOVA of the precision for the terminal deployed LoS sensor with the effects that are not statistically different removed. Table 41, Table 42, Table 43, Table 44 and Table 45 provide the coefficients for estimating the mean precision for the terminal LoS sensor. The constant coefficient for estimating the precision is 1.78 square feet.

Table 40. ANOVA of Terminal LoS Sensor Precision Without Three-Level Interaction

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Desk	15.0	14.11%	1	15.0	714	0
Number Cameras	38.9	36.47%	3	13.0	615	0
Number People	1.72	1.61%	2	0.86	40.9	0
Desk*Number People	0.338	0.32%	2	0.169	8.03	0.000335
Number Cameras*Number People	0.572	0.54%	6	0.0953	4.53	0.000142
Error	50.0	46.88%	2373	0.0211		
Total	107		2387			

Table 41. Presence of Desk Coefficients of Terminal LoS Sensor Mean Precision

Desk	Coefficient
With	-0.0793
Without	0.0793

Table 42. Number of Cameras Coefficients of Terminal LoS Sensor Mean Precision

Number Cameras	Coefficient
0	0.276
1	-0.00603
2	-0.104
3	-0.166

Table 43. Number of People Coefficients of Terminal LoS Sensor Mean Precision

Number People	Coefficient
30	-0.0415
60	0.00870
90	0.0328

Table 44. Number of People with Presence of Desk Coefficients of Terminal LoS Sensor Mean Precision

Number People	Desk	Coefficient
30	With	0.00216
	Without	0.0134
60	With	-0.0156
	Without	-0.00216
90	With	-0.0134
	Without	0.0156

Table 45. Number of People with Number of Cameras Coefficients of Terminal LoS Sensor Mean Precision

Number People	Number Cameras	Coefficient
30	0	-0.0481
	1	0.0234
	2	0.0246
	3	0.00384
60	0	-0.00367
	1	-0.000174
	2	0.0186
	3	-0.009487
90	0	-0.009124
	1	0.0256
	2	-0.0103
	3	-0.0153

The equation to estimate the mean precision for the terminal LoS sensor is:

$$\begin{aligned} \text{Terminal LoS Precision} = {} & 1.78 \\ & + \textit{Presence of Desk Coefficient} \\ & + \textit{Number People Coefficient} \\ & + \textit{Number Cameras Coefficient} \\ & + \textit{Number People with Presence of Desk Coefficient} \\ & + \textit{Number People with Number Cameras Coefficient} \end{aligned} \quad (6)$$

Table 46 and Figure 55 contain the mean and 95% confidence interval of the terminal LoS sensor given the presence or absence of the desk and Figure 56 shows a box plot of the distribution.

Table 46. Confidence Intervals of Terminal LoS Sensor Mean Precision Given Presence of Desk

Number People	Desk	Number Cameras	Mean	Error (95%)	Confidence Interval
30	With	0	1.87	0.0901	[1.96, 1.78]
		1	1.65	0.0219	[1.67, 1.63]
		2	1.58	0.0196	[1.60, 1.56]
		3	1.53	0.0315	[1.57, 1.50]
	Without	0	2.06	0.0735	[2.13, 1.99]
		1	1.82	0.0225	[1.84, 1.79]
		2	1.72	0.0148	[1.74, 1.71]
		3	1.66	0.0191	[1.68, 1.64]
60	With	0	2.05	0.127	[2.17, 1.92]
		1	1.70	0.0339	[1.74, 1.67]
		2	1.60	0.0132	[1.62, 1.59]
		3	1.55	0.0182	[1.57, 1.54]
	Without	0	2.13	0.0582	[2.18, 2.07]
		1	1.85	0.0146	[1.86, 1.83]
		2	1.74	0.0120	[1.76, 1.73]
		3	1.67	0.0151	[1.68, 1.65]
90	With	0	2.01	0.0792	[2.09, 1.93]
		1	1.70	0.0144	[1.72, 1.69]
		2	1.61	0.0120	[1.62, 1.59]
		3	1.55	0.0193	[1.56, 1.53]
	Without	0	2.21	0.0596	[2.27, 2.15]
		1	1.91	0.0146	[1.92, 1.90]
		2	1.79	0.0101	[1.80, 1.78]
		3	1.71	0.0109	[1.72, 1.71]

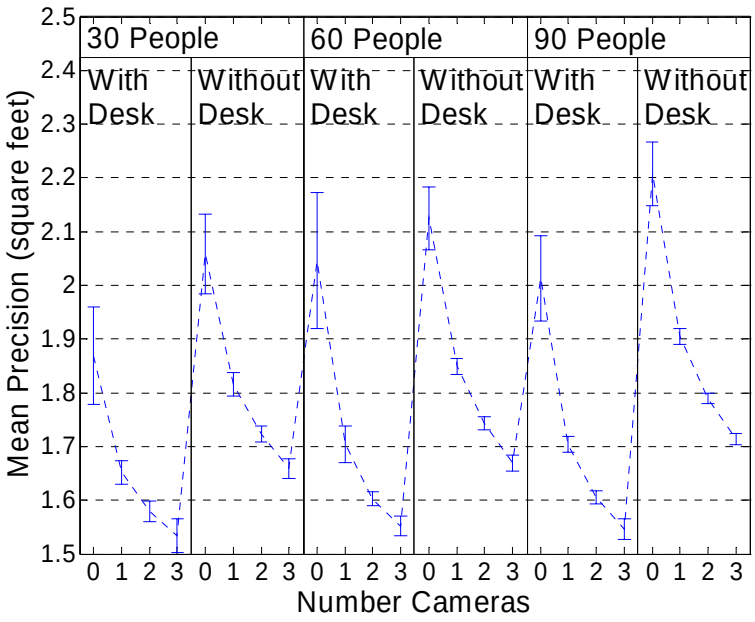


Figure 55. Confidence Interval of Terminal LoS Sensor Precision Given Number of People, Number of Cameras and Presence of Desk

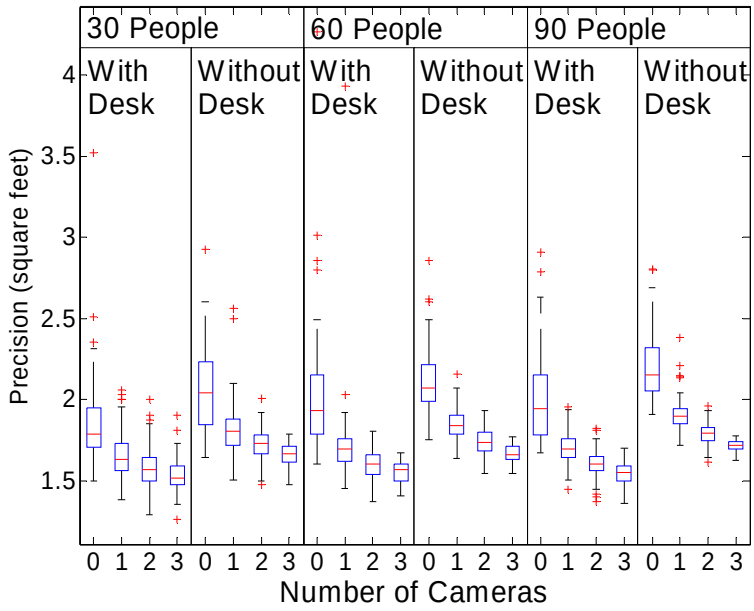


Figure 56. Distribution of Terminal LoS Sensor Precision Given Number of People, Number of Cameras and Presence of Desk

H. Three-Camera System Characteristics

The three-camera system is comprised of the three cameras used in the simulation, fused together. The results and analysis of the three-camera system is provided here as a comparison to the LoS sensor results. The two primary areas the three-camera sensor system is used are the corridor and the terminals. The corridor is monitored by one camera placed within the corridor in an adjacent ROI, and two other cameras placed on opposite sides of the same terminal adjacent to the ROI. The terminal is monitored by three cameras placed in the corridor. All cameras placed such that the focal point is towards the center of the ROI being monitored. The width of the camera angle ensures the entire ROI is within the field of view of the camera.

Because of the differences in the deployment of the three-camera system in the corridors and terminal, the two situations are examined together and separately with the presence or absence of the desk as a factor in the terminals.

1) Three-Camera System Hit Rate

Table 47 shows the results of ANOVA of the hit rate for simulation results for the three-camera system. The p-factors show that alone, the numbers of people are not a statistically significant in terms of variance.

Table 47. ANOVA of Three-Camera System Hit Rate

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Type Area	0.430	27.62%	2	0.215	87.2	0
Number People	0.00848	0.55%	2	0.00424	1.72	0.180
Type Area*Number People	0.0403	2.59%	4	0.0101	4.09	0.00290
Error	1.08	69.36%	438	0.00246		
Total	1.56		446			

a) Three-Camera System Hit Rate for Corridor

ANOVA of the precision for the corridor deployed three-camera system given a number of cameras is shown in Table 48. Table 49 and Figure 57 contain the mean and 95% confidence interval for each level of people and Figure 58 shows as a box plot of the distribution.

Table 48. ANOVA of Corridor Three-Camera System Hit Rate

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Number People	0.0324	4.99%	2	0.0162	3.837	0.0238
Error	0.616	95.01%	146	0.00422		
Total	0.648		148			

Table 49. Confidence Intervals of Corridor Three-Camera System Mean Hit Rate

Number Cameras	Mean	Error (95%)	Confidence Interval
0	0.920	0.0225	[0.942, 0.898]
1	0.928	0.0162	[0.944, 0.912]
2	0.894	0.0160	[0.910, 0.878]

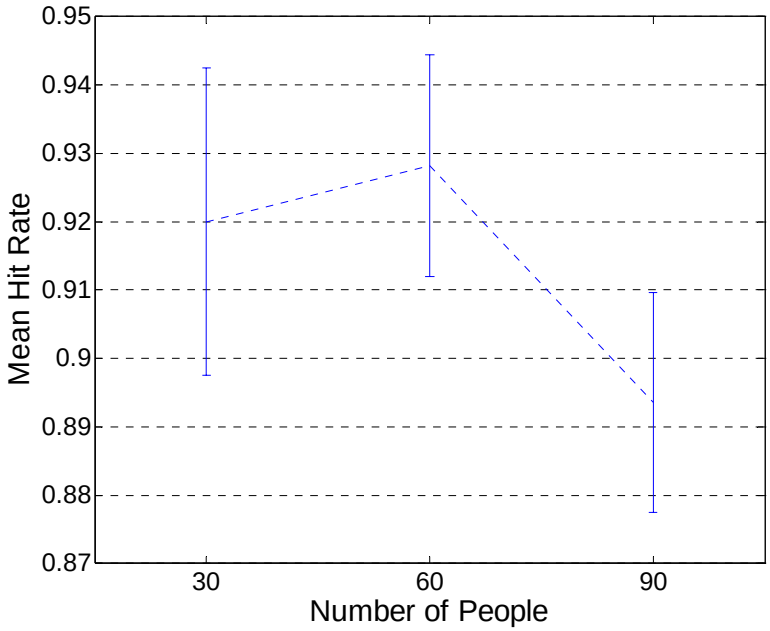


Figure 57. Confidence Interval of Corridor Three-Camera System Hit Rate

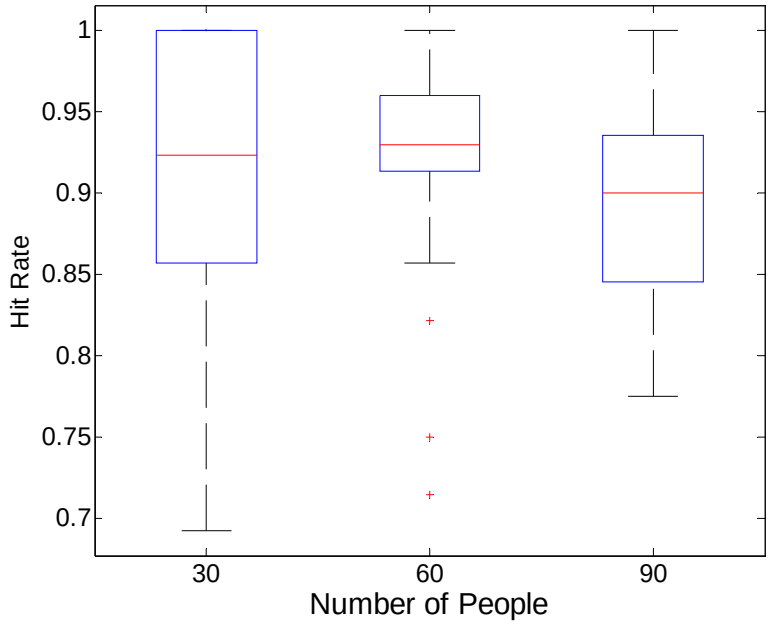


Figure 58. Distribution of Corridor Three-Camera System Hit Rate

b) Three-Camera System Hit Rate for Terminal

ANOVA of the precision for the corridor deployed three-camera system given a number of cameras is shown in Table 50. Table 51 and Figure 59 contain the mean and 95% confidence interval for each level of people and presence of desk and Figure 60 shows a box plot of the distribution.

Table 50. ANOVA of Terminal Three-Camera System Hit Rate

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Number People	0.00269	0.55%	2	0.00135	0.849	0.429
Desk	0.0134	2.72%	1	0.0134	8.46	0.00392
Number People*Desk	0.0137	2.78%	2	0.00685	4.32	0.0142
Error	0.464	93.95%	292	0.00159		
Total	0.493		297			

Table 51. Confidence Intervals of Terminal Three-Camera System Mean Hit Rate

Desk	Number People	Mean	Error (95%)	Confidence Interval
Without	30	0.981	0.0150	[0.996, 0.966]
	60	0.990	0.00655	[0.997, 0.984]
	90	0.985	0.00679	[0.992, 0.978]
With	30	0.982	0.0142	[0.997, 0.968]
	60	0.959	0.0125	[0.971, 0.947]
	90	0.975	0.00990	[0.984, 0.965]

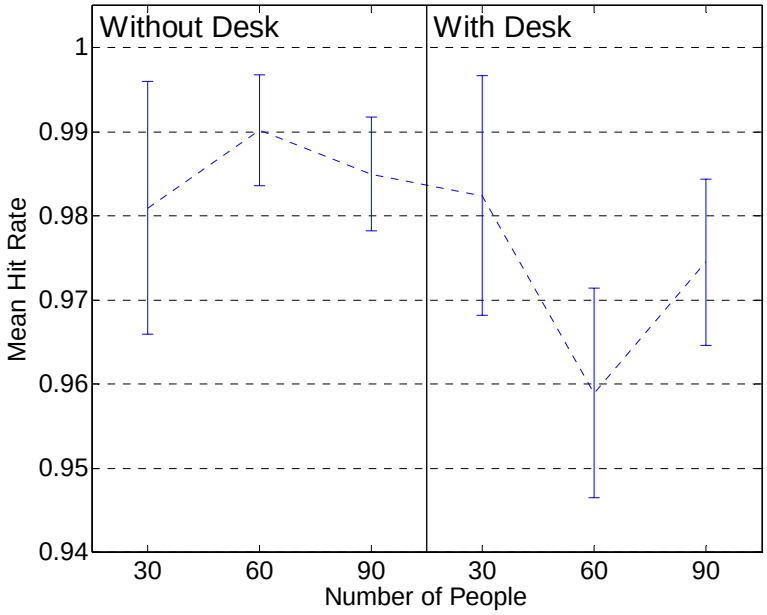


Figure 59. Confidence Interval of Terminal Three-Camera System Hit Rate

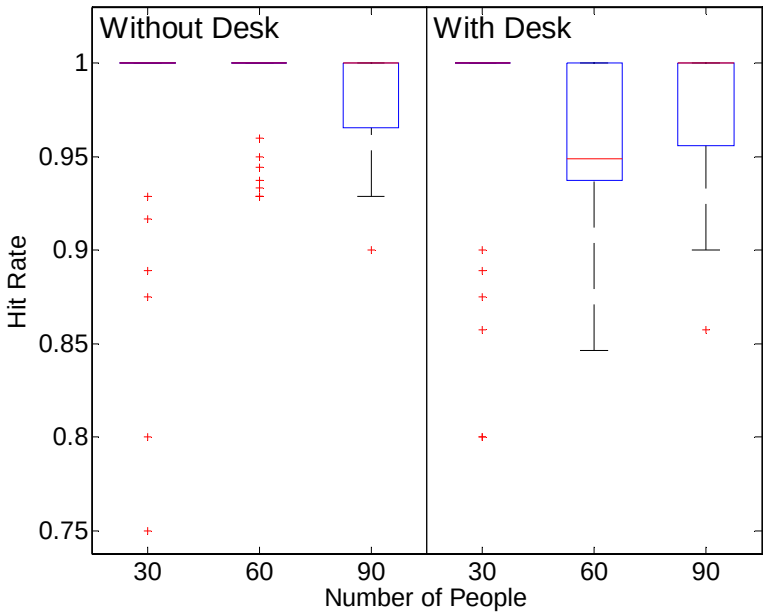


Figure 60. Distribution of Terminal Three-Camera System Hit Rate

2) Three-Camera System False Alarm Rate

Table 52 shows the results of ANOVA of the False Alarm rate for simulation results for the three-camera system. It can be seen from the p-factor that the contribution to the variance is statistically significant for all both factors considered independently or together.

Table 52. ANOVA of Three-Camera System False Alarm Rate

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Type Area	0.430	27.62%	2	0.215	87.2	0
Number People	0.00848	0.55%	2	0.00424	1.72	0.180
Type Area*Number People	0.0403	2.59%	4	0.0101	4.09	0.00290
Error	1.08	69.36%	438	0.00246		
Total	1.56		446			

a) Three-Camera False Alarm Rate for Corridor

ANOVA of the precision for the corridor deployed three-camera system given a number of cameras is shown in Table 53. Table 54 and Figure 61 contain the mean and 95% confidence interval for each level of people and Figure 62 shows box plot of the distribution.

Table 53. ANOVA of Corridor Three-Camera System False Alarm Rate

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Number People	88718	86.38%	2	44359	463	0
Error	13988	13.62%	146	95.8		
Total	102710		148			

Table 54. Confidence Intervals of Corridor Three-Camera System Mean False Alarm Rate

Number People	Mean	Error (95%)	Confidence Interval
30	4.92	1.02	[5.94, 3.90]
60	29.2	2.46	[31.7, 26.7]
90	64.5	4.08	[68.6, 60.4]

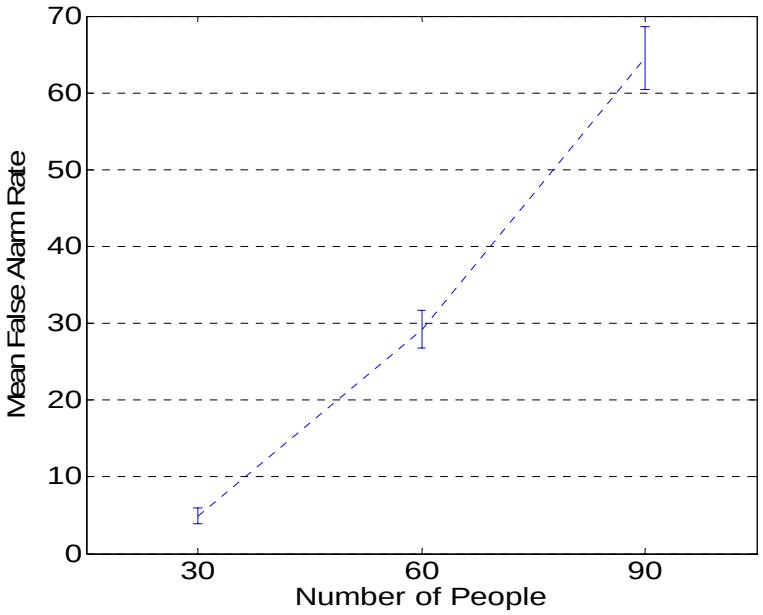


Figure 61. Confidence Interval of Corridor Three-Camera System False Alarm Rate

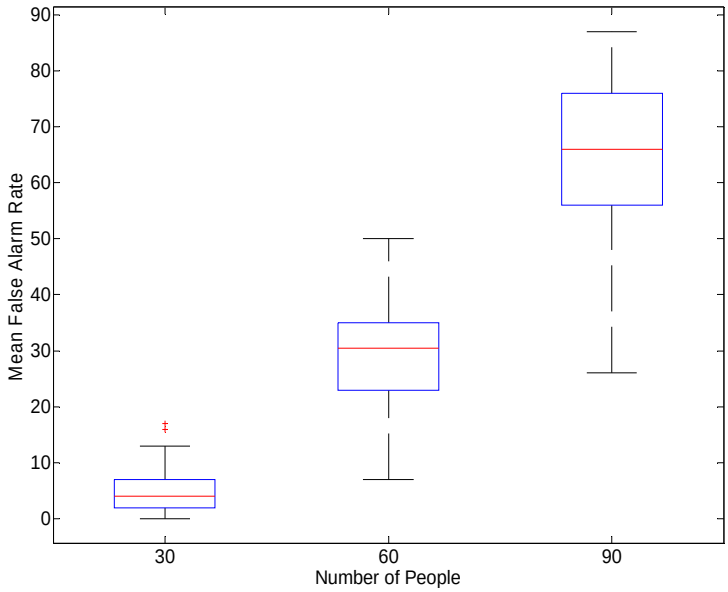


Figure 62. Distribution of Corridor Three-Camera System False Alarm Rate

b) Three-Camera System False Alarm Rate for Terminal

ANOVA of the precision for the corridor deployed three-camera system given a number of cameras is shown in Table 55. Table 56 and Figure 63 contain the mean and 95% confidence interval for each level of people and presence of desk and Figure 64 shows a box plot of the distribution.

Table 55. ANOVA of Terminal Three-Camera System False Alarm Rate

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Number People	10847	73.63%	2	5423	409	0
Desk	9.91	0.07%	1	9.91	0.748	0.388
Number People*Desk	5.12	0.03%	2	2.56	0.193	0.824
Error	3870	26.27%	292	13.3		
Total	14731		297			

Table 56. Confidence Intervals of Terminal Three-Camera System Mean False Alarm Rate

Desk	Number People	Mean	Error (95%)	Confidence Interval
Without	30	1.76	0.515	[2.28, 1.24]
	60	7.08	1.12	[8.20, 5.96]
	90	16.7	1.39	[18.1, 15.3]
With	30	1.68	0.435	[2.16, 1.24]
	60	6.78	0.811	[7.59, 5.97]
	90	16.0	1.50	[17.5, 14.5]

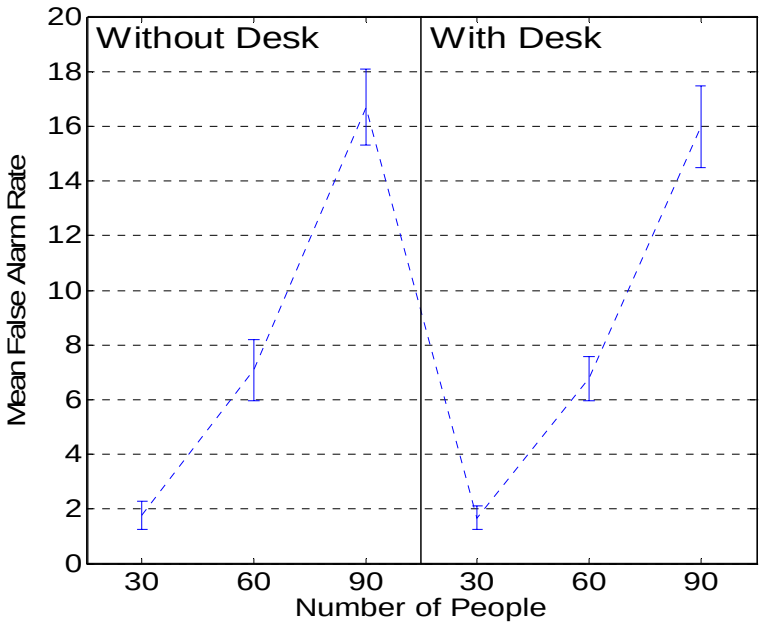


Figure 63. Confidence Interval of Terminal Three-Camera System False Alarm Rate

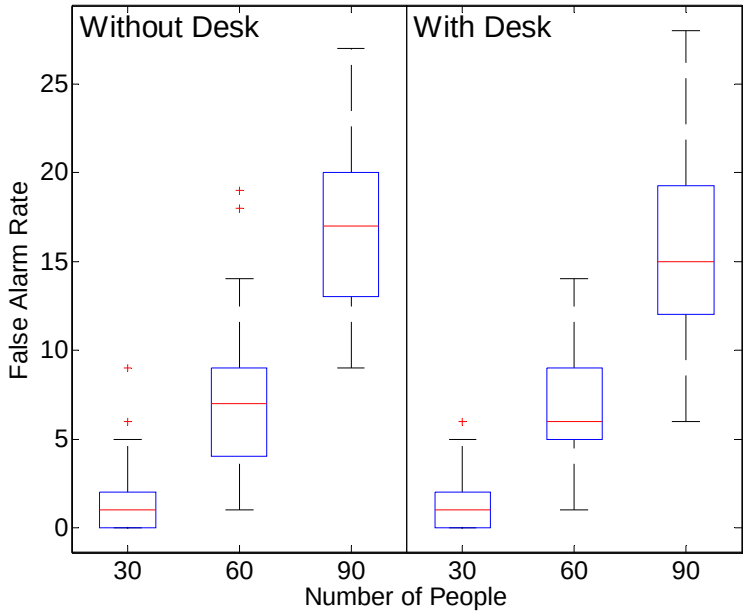


Figure 64. Distribution of Terminal Three-Camera System False Alarm Rate

3) Three-Camera System Precision

Table 57 shows the results of ANOVA of the Precision for simulation results for the three-camera system. It can be seen from the p-factor that the contribution to the variance is statistically significant for all both factors considered independently or together.

Table 57. ANOVA of Three-Camera System Precision

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Number People	166	21.93%	2	83.0	136	0
Desk	316	41.76%	2	158	260	0
Number People*Desk	7.47	0.99%	4	1.87	3.06	0.0165
Error	267	35.24%	438	0.609		
Total	757		446			

a) Three-Camera System Precision for Corridor

ANOVA of the precision for the corridor deployed three-camera system given a number of cameras is shown in Table 58.

Table 59 presents the mean and confidence interval for each level of people, and Figure 66 shows a comparison of the 95% confidence interval as well as a box plot of the distribution of the hit rate given the number of people.

Table 58. ANOVA of Corridor Three-Camera System Precision

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Number People	84.2	62.49%	2	42.1	122	0
Error	50.5	37.51%	146	0.346		
Total	135		148			

Table 59. Confidence Intervals of Corridor Three-Camera System Mean Precision

Number People	Mean	Error (95%)	Confidence Interval
30	4.40	0.121	[4.52, 4.28]
60	4.99	0.158	[5.15, 4.83]
90	6.21	0.213	[6.42, 6.00]

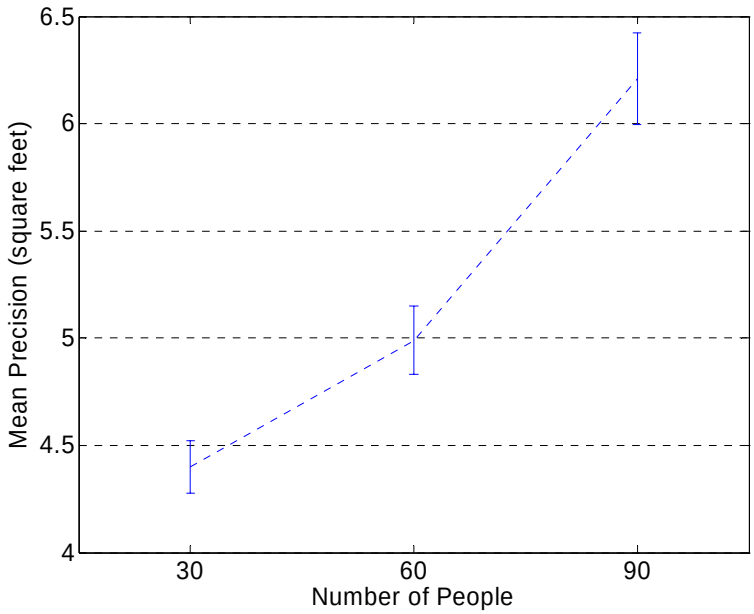


Figure 65. Confidence Interval of Corridor Three-Camera System Precision

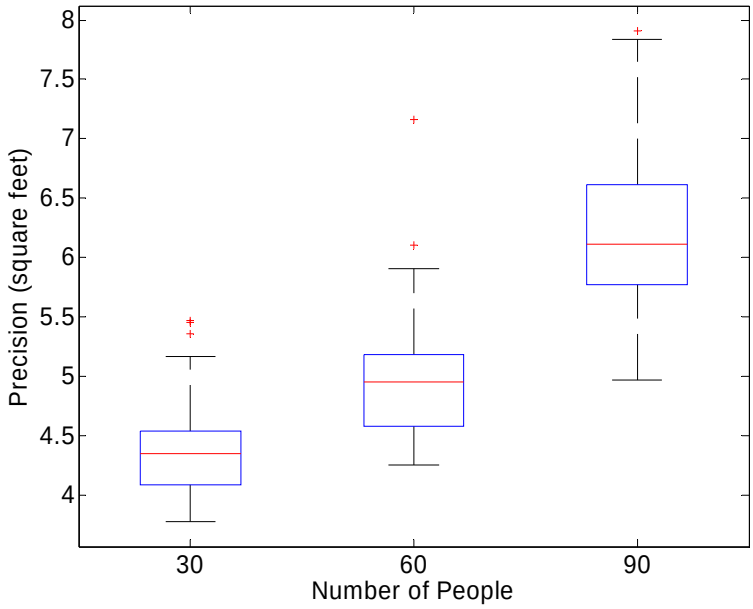


Figure 66. Distribution of Corridor Three-Camera System Precision

b) Three-Camera System Precision for Terminal

ANOVA of the precision for the corridor deployed three-camera system given a number of cameras is shown in Table 60. Table 61 and Figure 67 contain the mean and 95% confidence interval for each level of people and presence of desk and Figure 68 shows a box plot of the distribution.

Table 60. ANOVA of Terminal Three-Camera System Precision

Source	Sum Sq.	% Var	d.f.	Mean Sq.	F	Prob>F
Number People	86.5	26.73%	2	43.2	58.4	0
Desk	18.0	5.55%	1	18.0	24.3	0.00000141
Number People*Desk	2.84	0.88%	2	1.42	1.91	0.149
Error	216	66.87%	292	0.741		
Total	323		297			

Table 61. Confidence Intervals of Terminal Three-Camera System Mean Precision

Desk	Number People	Mean	Error (95%)	Confidence Interval
Without	30	6.59	0.227	[6.82, 6.37]
	60	6.94	0.211	[7.15, 6.73]
	90	8.00	0.376	[8.37, 7.62]
With	30	6.07	0.193	[6.26, 5.88]
	60	6.70	0.236	[6.94, 6.47]
	90	7.28	0.182	[7.47, 7.10]

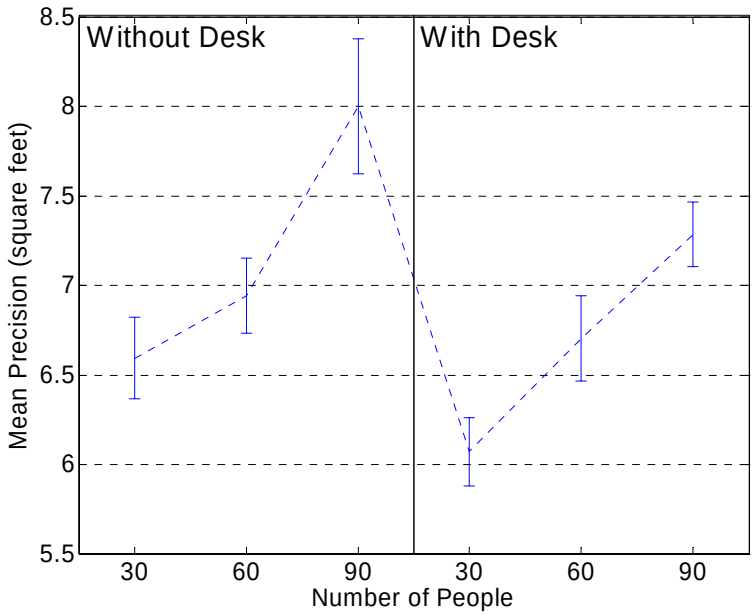


Figure 67. Confidence Interval of Terminal Three-Camera System Precision

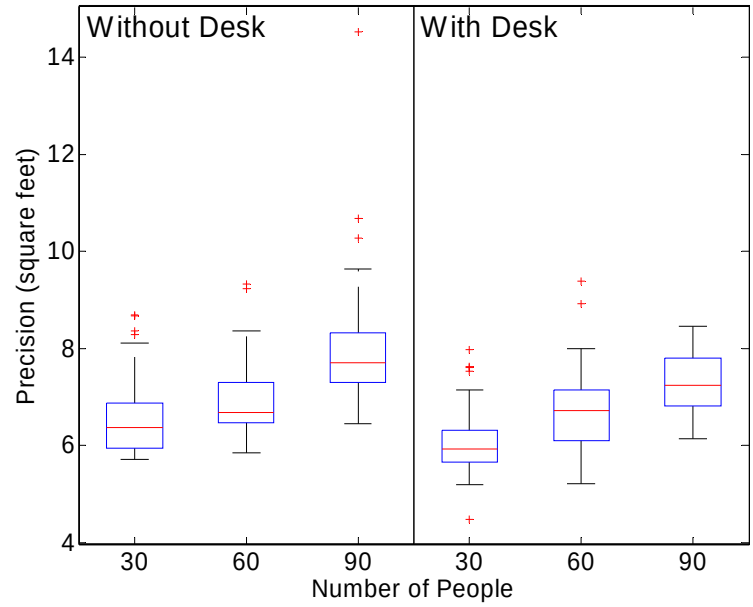


Figure 68. Distribution of Terminal Three-Camera System Precision

I. Summary

This chapter presented the results of the simulation for the configurations using the LoS sensor and configurations using three cameras without the LoS sensor. ANOVA is shown for the results, and results for which factors and interactions of factors are shown to contribute to the variance are examined for mean value, confidence interval, and variance.

V. Conclusions and Recommendations

A. Chapter Overview

The purpose of this chapter is to present the conclusion of the research based on the simulation results. The effect of each factor on hit rate, false alarm rate, and precision are presented in concise form and the reason for the effects is explained. In addition, a description of configurations using the LoS sensor and the configuration of three cameras are compared for robustness as factor values change. Finally, opportunities for further research of the LoS sensor and fusion of the LoS sensor with video networks are offered.

B. Conclusions of Research

1) Hit Rate

The primary reasons a LoS sensor does not detect a person is that the person is so small that not enough lines are blocked at a grid square, the person is in a location the LoS is not positioned to detect, or other people cause an undetectable area to develop. Each person is represented as a circle with a diameter four times greater than the width of a grid square so people occupy an entire grid square. However, limitations on sensor placement and obstructions caused by the terminal desk and other people contribute to the miss rate for the LoS sensor.

A camera does not detect a person when they resemble the background. All background colors are rendered in shades of gray, and all people are rendered in colors other than gray. However, some colors are so similar to background colors they cannot be distinguished by the differencing technique. Figure 69 is an example of two people detected by the camera. The person on the right is partially detected, the person on the

left is detected completely. The images are simplistic, noise-free background and foreground objects compared to real-world images. Figure 70 shows a camera image from the simulation and a real-world camera image. The simulation image is a Portable Network Graphics (PNG) image, the real image is a Portable Gray Map (PGM) image. Both are of pedestrian traffic, but the real image is noticeably more noisy and complex though it contains fewer people.

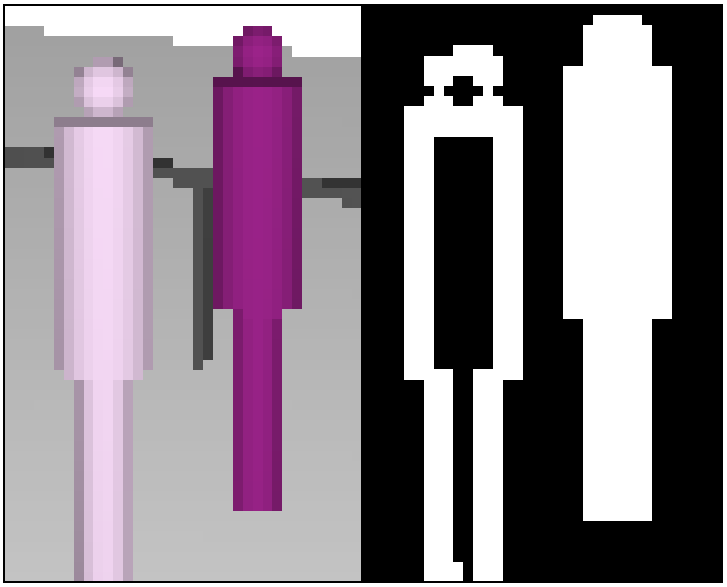


Figure 69. Example of Camera Image with Partial Detection

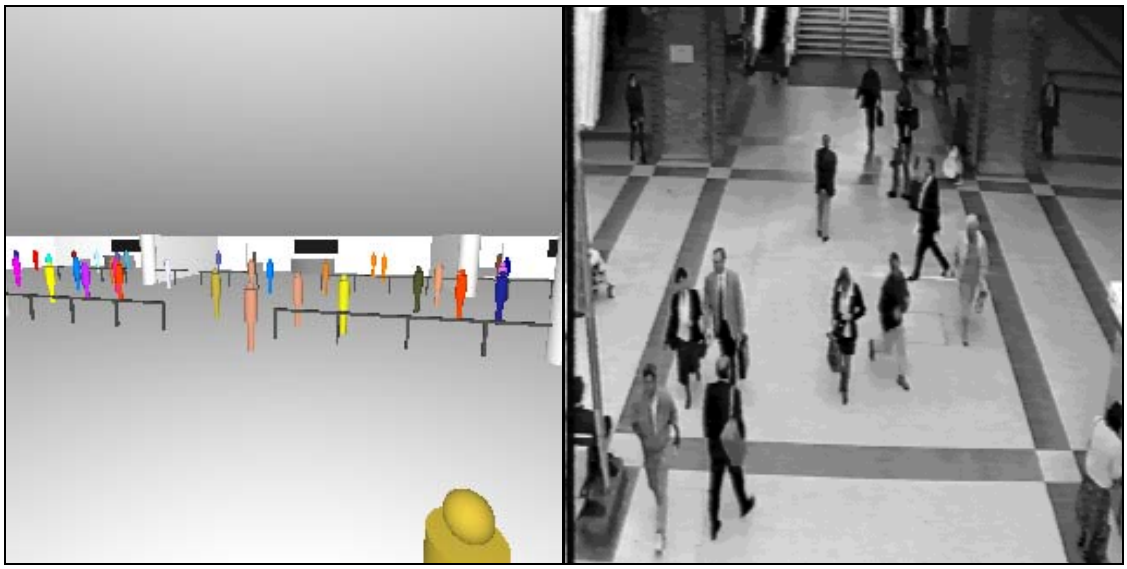


Figure 70. Example of Simulated Camera Image with Real-World Camera Image of Pedestrian Traffic

Another reason some people may not be detected by either a camera or the LoS sensor is they appear to be merged with a nearby person. The algorithm that counts the number of hits per scenario associates each connected component of the detected image with at most one person. If two people are so close together that they appear as a single larger person, the algorithm detects a single person.

The results of a corridor simulation in Figure 71 and Figure 72 show a LoS sensor without any cameras. This configuration provides a higher mean hit rate than a three-camera system. Figure 73 and Figure 75 show the mean LoS sensor hit rate for a terminal. It is less than that of the three-camera system. The reason for the higher performance of the LoS sensor operating independently in the corridor is due to interference of the cameras by activity in adjacent ROIs. A camera is placed in one ROI

to observe another ROI, but also observes a ROI beyond that. Therefore, people moving immediately in front of or behind the ROI the camera is observing may result in added person location hypotheses. These extra hypotheses may induce people to merge in the camera sensor. This phenomenon is occurs more often in a corridor than in a terminal, as a corridor is open on all sides while a terminal is open on only one side. Indeed, the mean hit rate for the camera operating in the terminal is statistically higher than the corridor for most scenarios. The LoS sensor operating in a corridor ROI is also affected by activity in an adjacent corridor ROI, as sensors in one corridor ROI can operate with the lights of an adjacent LoS sensor in a neighboring corridor ROI. However, distribution of the LoS sensor light rays prevents the interference from a neighboring ROI from causing as many false hypotheses that would normally result in people within the ROI merging.

Figure 71 shows that the hit rate decreases as more cameras are fused with the LoS sensor. The decline in LoS hit rate performance as cameras are added is due to the AND fusing algorithm. If one sensor hypothesizes a person exists at a given location, data from another sensor may either confirm or eliminate that hypothesis. However, if one sensor does not hypothesize that a person exists at a given location, then data fused from another sensor using an AND algorithm cannot create a hypothesis of a person existing at that location. Therefore, increasing the number of fused sensors using an AND algorithm either decrease or leave unchanged the number of hits for a given scenario, but cannot increase the number of hits for the same scenario.

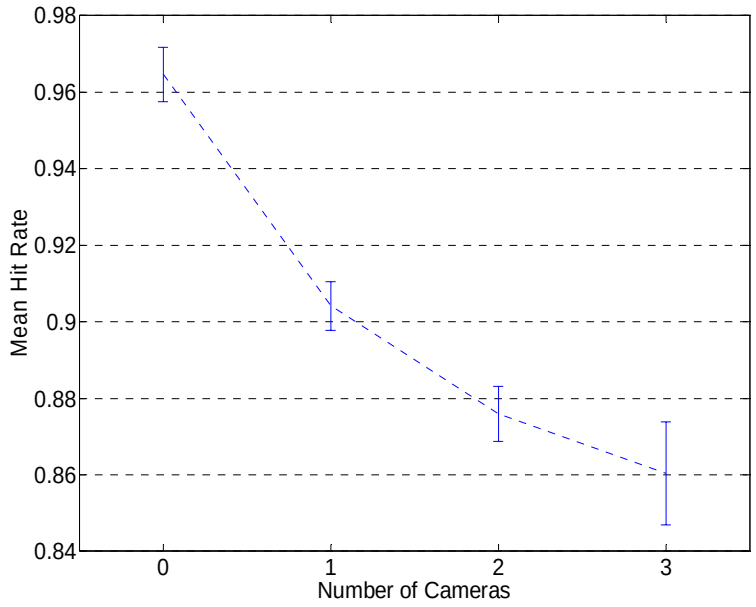


Figure 71. LoS Sensor Corridor Hit Rate

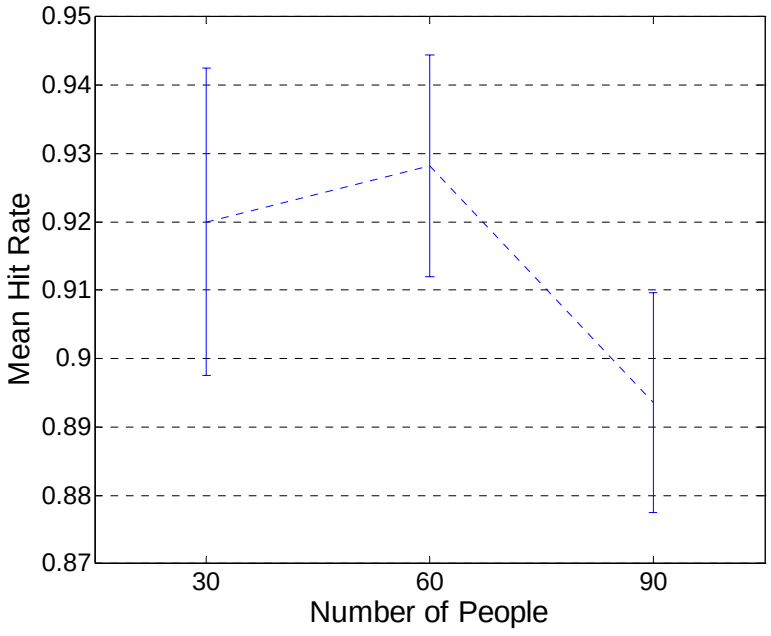


Figure 72. Three-Camera System Corridor Hit Rate

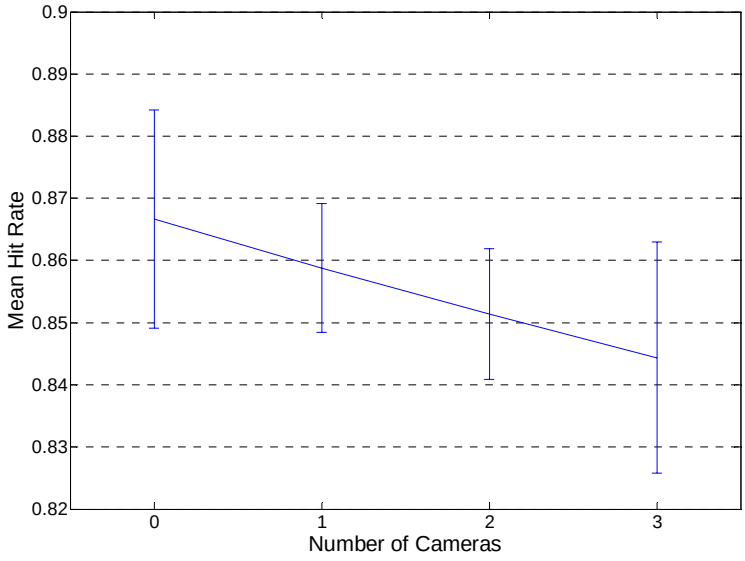


Figure 73. LoS Sensor Terminal Hit Rate Given Number of Cameras

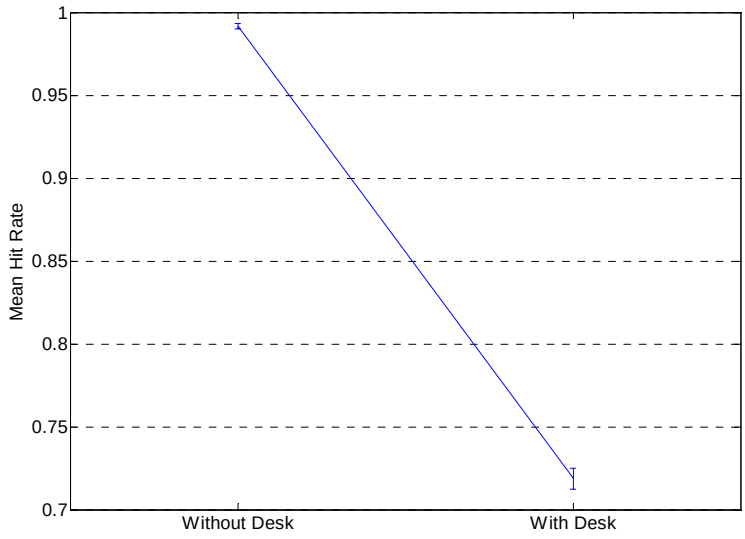


Figure 74. LoS Sensor Terminal Hit Rate Given Presence of Desk

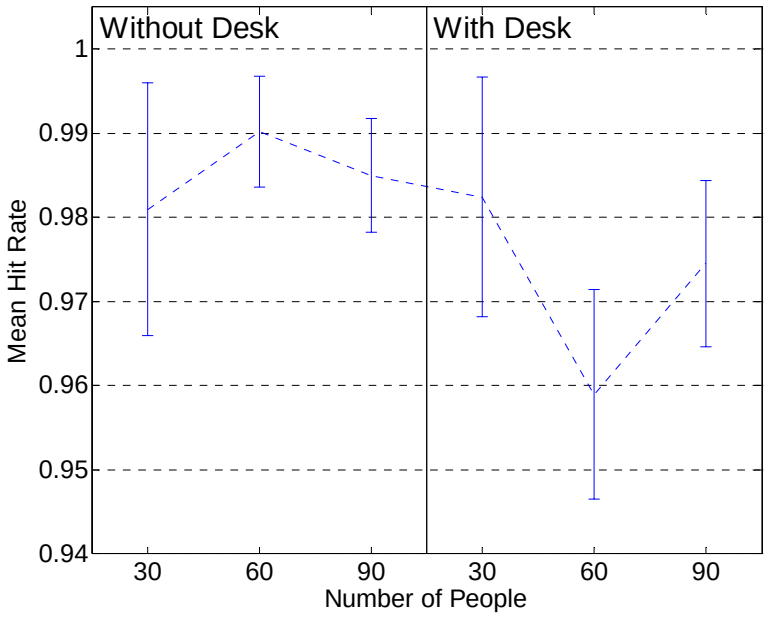


Figure 75. Three-Camera System Terminal Hit Rate

Analysis of Variance (ANOVA) showed the number of people in the simulation does not result in a statistically significant difference in the LoS sensor hit rate at the 95% confidence interval for the 50 replications performed. The same is true for the three-camera system observing the terminal, but the not for the three-camera system observing the corridor. Again, this is due to the higher susceptibility of the camera system to interference from adjacent ROIs in the corridor than the terminal. In the terminal, the camera system shows a statistically significant decrease with 60 people in the simulation when a desk is present.

The decrease in camera performance with the desk is because to the additional image complexity produced by the desk results in a higher likelihood of partial person detection. The LoS sensor shows a more dramatic decrease in mean hit rate due to the

desk. This can be attributed to the light rays blocked by the desk. A camera is able to detect people in front of the desk, but for the LoS sensor, any background obstruction along the length of the light ray renders that light ray unusable. This loss of light rays results in "dead" areas, such as the sides of the desk. It is noteworthy that the desk and people both block LoS sensor light rays, but only the desk results in a statistically significant difference in the mean hit rate. The difference is the desk obstructs a contiguous grouping of light rays, whereas adding people obstructs randomly distributed light rays.

For mean hit rate, the three-camera system is more robust with large obstructions and has fewer undetectable areas than the LoS sensor. The mean hit rate for the LoS sensor appears more robust to changes in the number of people monitored and is less affected by events occurring in adjacent areas.

2) False Alarm Rate

A false alarm occurs in the LoS if the percentage of blocked lights that cross a grid square is equal to the threshold value, even though no person occupies the grid square. This occurs when light rays intersecting a grid square are actually blocked by people existing elsewhere in the ROI. This phenomenon is more likely to occur in grid squares with fewer light rays intersecting it, such as the terminal entrances. A camera produces false alarms for a similar reason as the LoS sensor, except detection is due to the light rays hitting a non-background person rather than being blocked or unblocked.

Figures 75 through 78 shows the mean false alarm rate is lower for the LoS sensor than for the three-camera system for any given combination of factor levels. Increasing the number of people results in an increase in the false alarm rate for both the LoS sensor

and three-camera system. However, with 60 or 90 people in the simulation, the mean false alarm rate of the three-camera system is an order of magnitude greater than the LoS sensor system for both corridor and terminals.

One cause of the reduced false alarm rate of the LoS sensor is that the light rays used to remove hypotheses are more widely distributed than are the light rays observed by the cameras. Another reason for the higher performance of the LoS sensor is due to interference of the cameras by activity in adjacent ROIs. A camera placed to observe one ROI, also observes a ROI beyond that. Therefore, people moving immediately in front of or behind the ROI the camera is observing may result in false hypotheses of people location. This phenomenon is expected to occur more often in a corridor than in a terminal, as a corridor is open on all sides and a terminal is open on only one side. Indeed, the mean false alarm rate for the camera operating in the terminal is statistically higher than the corridor. The LoS sensor operating in a corridor ROI is also affected by activity in an adjacent corridor ROI, as sensor in one corridor ROI operates with the lights of an adjacent LoS sensor in a neighboring corridor ROI. However, distribution of the LoS sensor light rays prevents the interference from a neighboring ROI from causing as many false hypotheses.

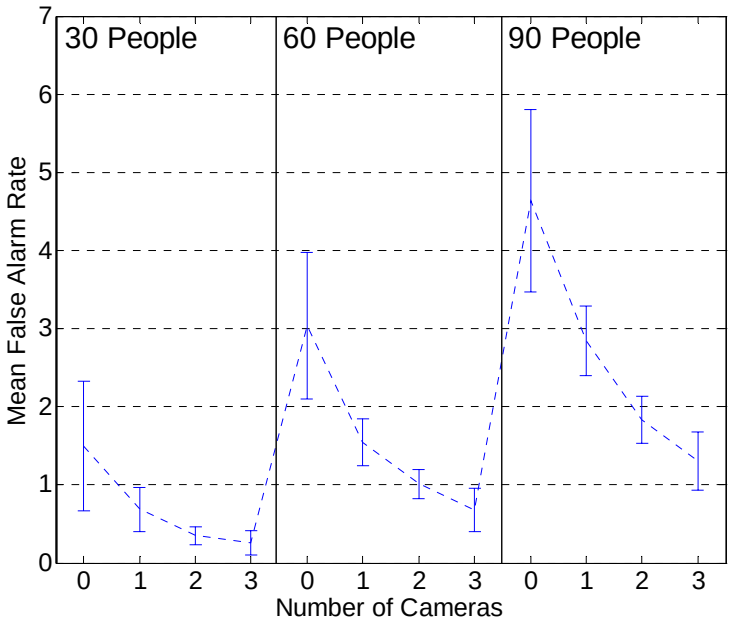


Figure 76. Corridor LoS Sensor False Alarm Rate

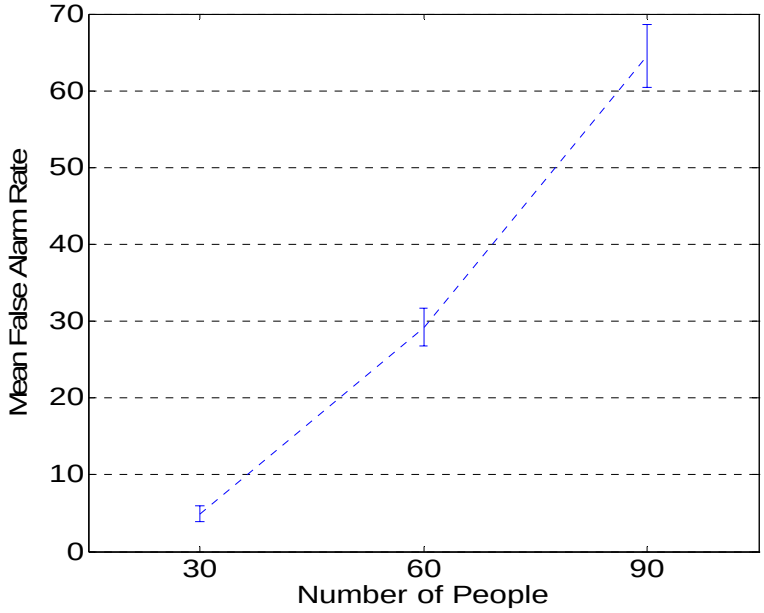


Figure 77. Corridor Three-Camera System False Alarm Rate

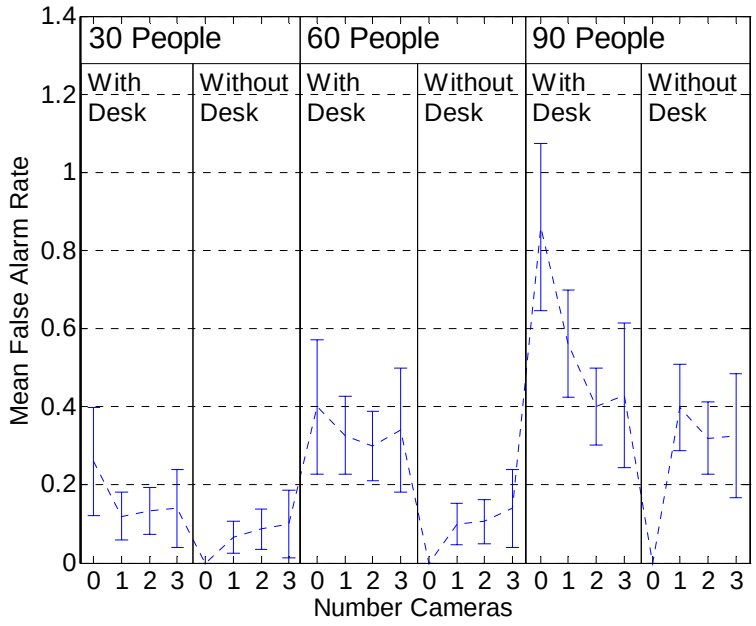


Figure 78. Terminal LoS Sensor False Alarm Rate

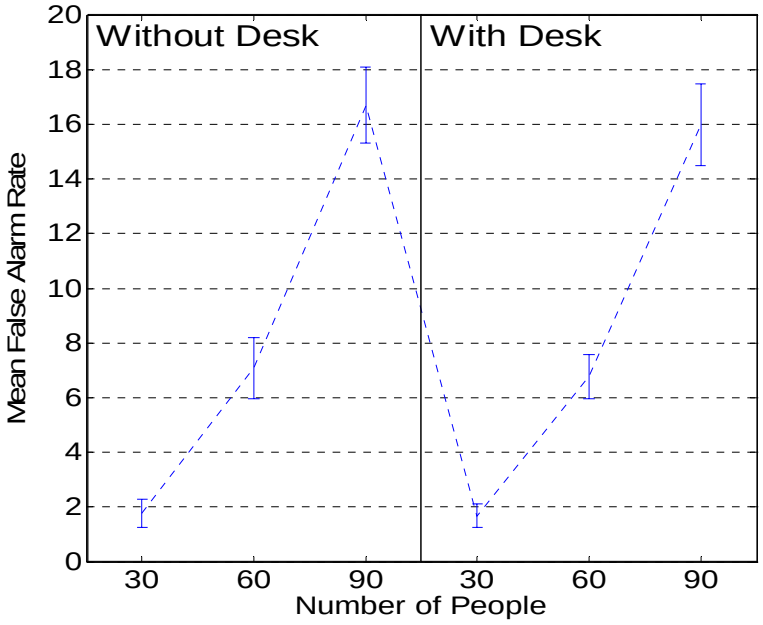


Figure 79. Terminal Three-Camera System False Alarm Rate

It is noteworthy that in terminals without desks and no cameras, the mean false alarm rate is very near zero but increases when cameras are added. One reason for the increase is the cameras remove portions of correct hypotheses about the location of a person without removing every portion of a correct hypothesis. An example would be an area that the LoS hypothesizes contains a person, whereas the true location of the person is on the left side but within the indicated area. This is a hit. However, the fusion of this hypothetical data with other sensor data removes only the left side of the hypothesized area. The remaining area no longer contains the person, and is now determined to be a false alarm rather than a hit.

Another possibility, similar to the above example, is the camera data splits a hypothesized region into two components. In the previous example instead of removing the entire left side of the hypothesized region, only the middle portion is removed leaving a left and right side. The left side of the hypothesized region still contains the person and is determined to be a hit. However, the remaining disconnected right side does not contain the person and is now determined to be a false alarm.

The mean false alarm rate for the LoS sensor is lower than the three-camera system for any given combination of factor levels, often by an order of magnitude. The false alarm rate is lowered further by fusing the LoS with an increasing number of cameras. Increasing the number of people in the scenario increases the mean false alarm rate for both LoS sensor and the three-camera system. The presence of large obstacles statistically increases the mean false alarm rate of the LoS sensor when used alone, but does not statistically increase the mean false alarm rate when fused with three cameras.

3) Precision

Precision is the area corresponding to the number of grid squares involved in the correct determination of location. Because the radius of each person is 0.557 feet, the ideal precision for the simulation would be the area of the circle defining a person, 0.975 square feet. All sensor data is rectified to three-inch-by-three-inch grid squares within the ROIs. The minimum precision resolution for any sensor in this simulation is then 0.0625 square feet.

The precision of the LoS sensor can deteriorate when there are not enough light rays to remove false hypotheses which occurs for example when a person approaches the extremity of the corridor. In these areas, many vertical light rays determine the horizontal hypotheses of people locations, but there are few diagonal light rays to remove hypotheses in the vertical direction. This results in a loss of precision for the LoS sensor, primarily in the vertical direction.

The precision of the three-camera system deteriorates when lower resolution cameras are used. As the camera resolution increases, more pixels are available to more accurately define the outer edges of the detected person. For example, Figure 81 shows two cameras detecting the location of people. The precision of the camera as the resolution increases is limited to a four-sided polygon containing the detected person.

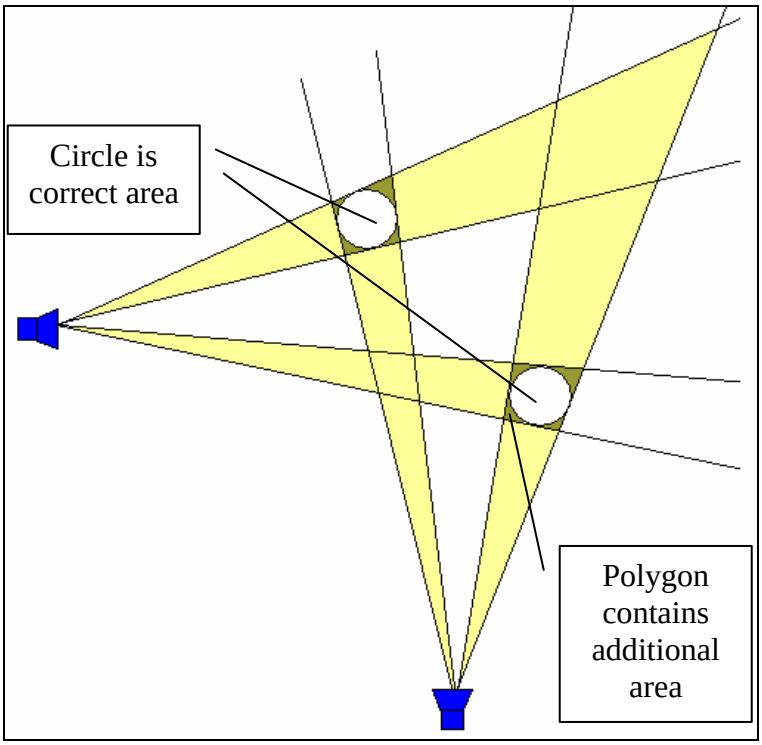


Figure 80

Figure 81. Example of Two Cameras Detecting Objects

Figure 82 shows a similar system with three cameras detecting the location of people. The precision of the three-camera system improves upon the two-camera system. As the resolution increases, the precision for the three-camera system is limited by a six-sided polygon containing the detected person.

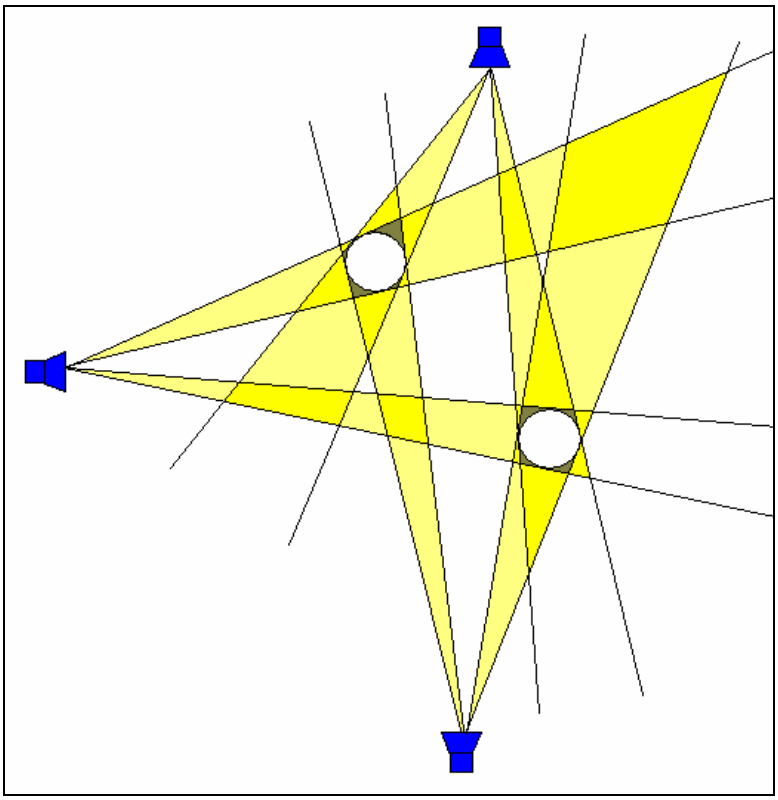


Figure 82. Example of Three Cameras Detecting Objects

In Figure 84, the black circle is the person detected by the LoS sensor, represented by black bars. Black lines show tangents drawn from the extremities of the LoS sensor to the outer edge of the detected person. The dark gray areas in Figure 84 contain no LoS sensor light-rays, but the dark gray areas are also not occupied by the person. Figure 84 shows an detected person appears that elongated towards the LoS sensor light sources and sensors, with limits defined by tangents drawn from the extremities of the LoS sensor to the outer edges of the person.

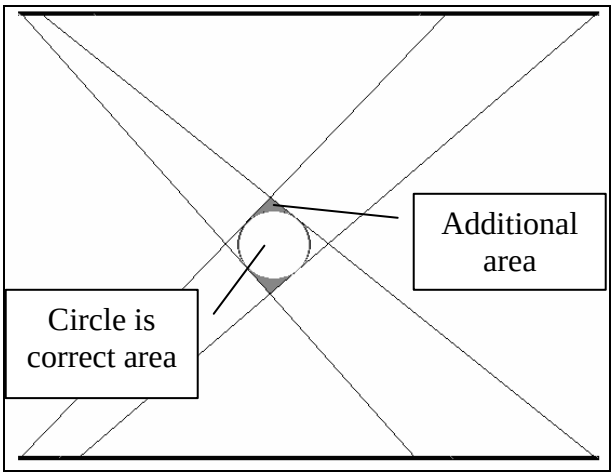


Figure 83

Figure 84. Example of LoS Sensor Detecting Object

Figures 84 thru 88 shows that the LoS sensor provides consistently better precision than the three-camera system at all factor levels. Figure 85 and Figure 86 show that in the corridor, both the LoS sensor and the thee-camera system lose precision as the number of people increases from 30 to 60. However, as the number of people continues to increase from 60 to 90, the three-camera system continues to lose precision while the LoS sensor's precision is not statistically different. Figure 87 indicates that although the LoS sensor's precision is superior to that of the camera system, it is further improved by the addition of one or two cameras. However, there is not a statistical difference between the mean precision of the LoS sensor with two or three cameras. In the terminal, the three-camera system loses precision just as the terminal three-camera system does, but the LoS sensor for the terminal shows no statistical difference in the mean of the precision as the number of people increases.

In the terminal, it can be seen that the presence of a desk results in a loss of precision for both the three-camera system and LoS sensor. The cameras loss of precision is due to the additional image complexity of the desk. For the LoS sensor, the loss of precision is due to the reduction in available light-rays to remove false hypotheses.

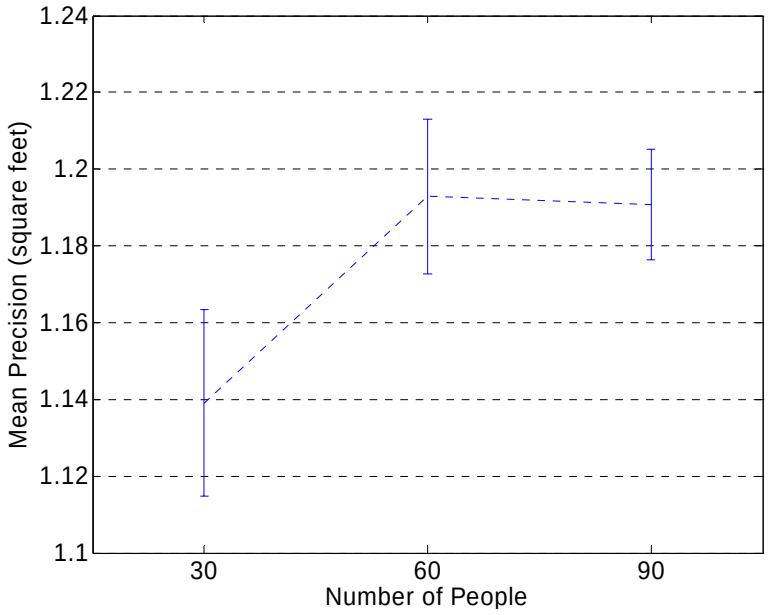


Figure 85. Corridor LoS Sensor Precision Given Number of People

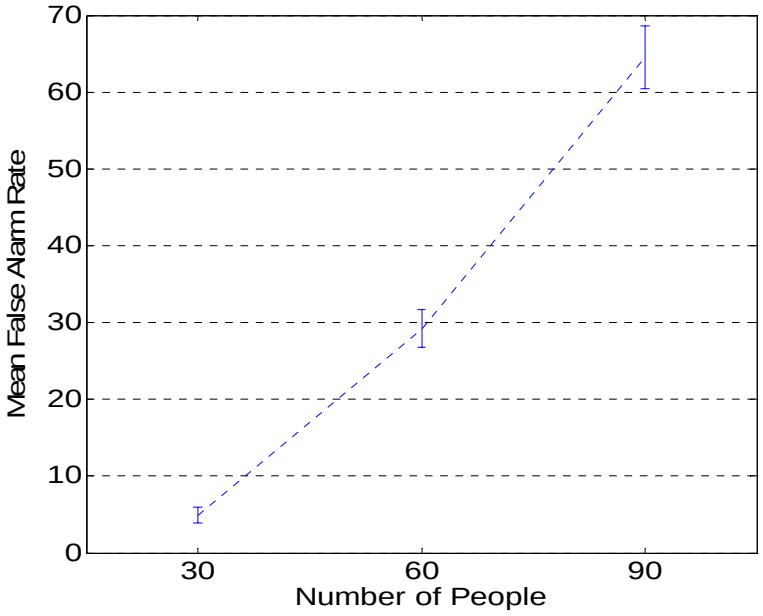


Figure 86. Corridor Three-Camera System Precision

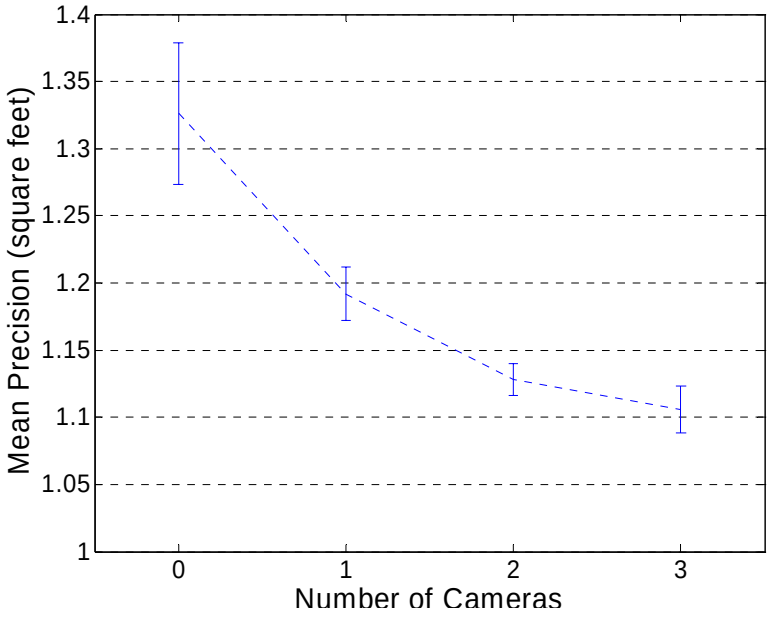


Figure 87. Corridor LoS Sensor Precision Given Number of Cameras

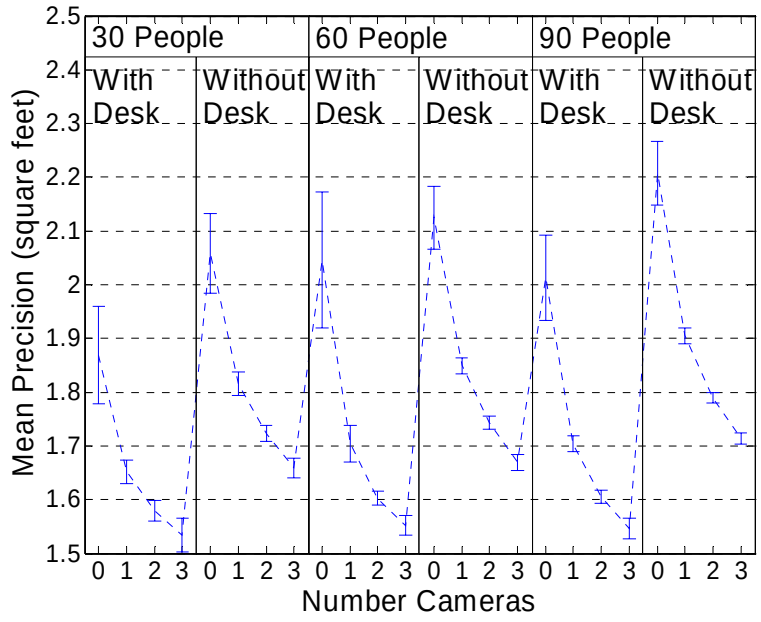


Figure 88. Terminal LoS Sensor Precision

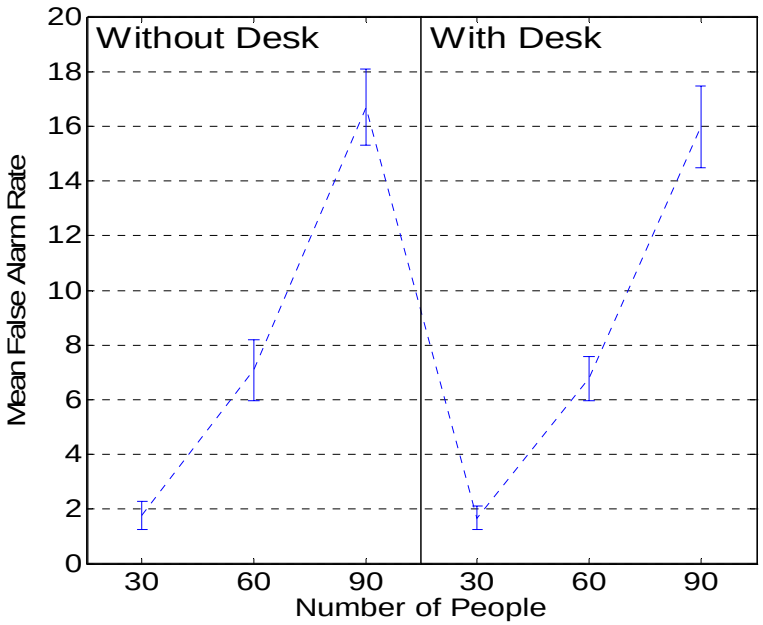


Figure 89. Terminal Three-Camera System Precision

C. Summary

On average, the simulated camera system achieves a higher hit rate than the LoS sensor, but the LoS sensor offers a better false alarm rate and precision. The higher mean hit rate of the three-camera system is because the simulation images are not as complex as are images from a real system. In addition, the LoS sensor cannot detect portions of an area being observed. The increased mean false alarm rate of the three-camera system over the LoS sensor is due to the distribution of LoS sensors. This accounts for part of the improved precision of the LoS sensor. However, the LoS images are smoothed before threshold application by a Gaussian filter whereas the camera images are not. This smoothing reduces the area occupied by detected people. In addition, the area people detected by cameras decrease as the image resolution increases. Loss of precision attributable to Gaussian smoothing and camera resolution does not alter the trends of the LoS sensor or three-camera system as the factor levels change.

Increasing the number of people resulted in little effect on the hit rate for the LoS sensor system and the three-camera system. The increase from 60 to 90 people does result in a statistically significant decrease in the mean hit rate of the LoS sensor. However, the mean hit rate of the LoS sensor for 30 people is not statistically different from the mean hit rate for 60 or 90 people. The mean hit rate of the three-camera system is not statistically different for any of the three people levels.

Increasing the number of people in the simulation increases the false alarm rate for the LoS sensor and the three-camera system. However, the false alarm rate does not increase as dramatically with the number of people as does the three-camera system.

Increasing the number of people also increases the area each target is hypothesized to occupy for the LoS sensor and the three-camera system. For the LoS sensor, increasing the number of people from 60 to 90 results in no significant change in precision, whereas an increase from 30 people to 60 people does. Whereas for the three-camera system, the change in the mean precision as the number of people increases from 60 to 90 is greater than the change in the mean precision as the number of people is increases from 30 to 60. This suggests that a larger number of people in a given area does not affect the precision of the Los sensor significantly.

For the terminal areas, Figure 74, Figure 75, Figure 78, and Figure 79 show that the mean hit-rate and mean false alarm rate of the LoS sensor is more affected by the presence of a desk than is the mean hit rate and mean false alarm rate for the three-camera system. However, no people ever travel behind the desk from the perspective of a camera. Thus, inclusion or removal of the desk only changes the background image for the three-camera system, and never the source of foreground occlusion of a person. Because sensing for the LoS sensor is more distributed than the camera system, foreground and background do not as readily apply. For the LoS sensor, adding the desk results in loss of light source-sensor pairs participating in the sensing activity. Figure 90 and Figure 91 show the differences in the number of LoS sensor light rays passing through each grid square. Each square of these figures corresponds to a grid square observed by the LoS sensor, and the brightness of the square represents the number of light source to sensor segments intersecting it. Figure 90 is a LoS sensor without a desk, and Figure 91 is a LoS sensor with a desk. An image of the desk is added to Figure 91 for reference. The effect of the desk is to reduce the number of light rays for each grid

square below tangent lines from extremities of the LoS sensor to the outer edges of the desk, as shown with the darkening of the lower part of Figure 91.

The presence of the desk adversely affects the precision of the LoS sensor system and three-camera system for most scenarios. The decline in precision for the LoS is again due to the reduction in light source-sensor pairs available for detection. The decrease in precision for the camera system is due to the added complexity of the background image. However, because the desk is always in the background, the people are still detected.

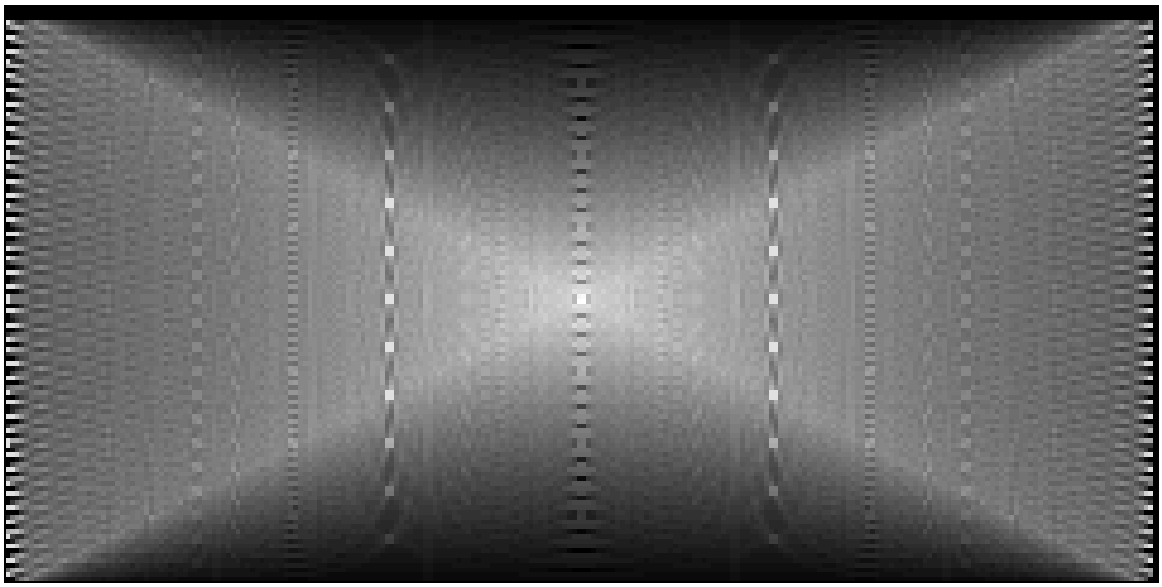


Figure 90. Comparison Array of LoS Sensor in Terminal without Desk

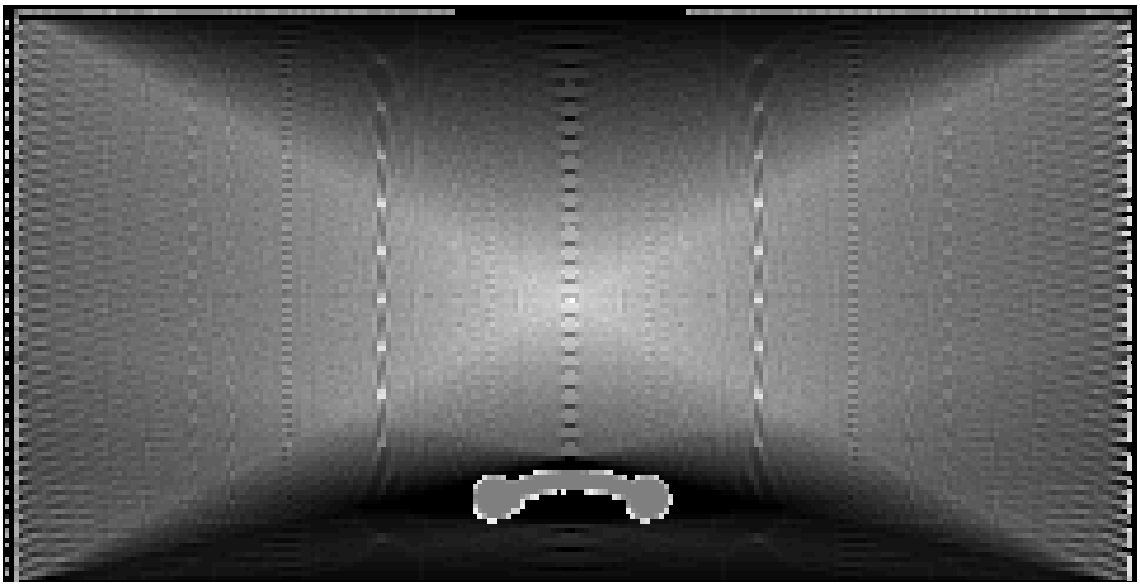


Figure 91. Comparison Array of LoS Sensor in Terminal with Desk Superimposed On Image

The LoS sensor was tested fused with zero to three cameras. As the number of cameras increases the mean hit rate declines. This is expected because the results are ANDed together. Any person detected by a LoS sensor that is not also detected by a camera is not present in the fused data. Therefore, as more cameras are available to remove people from the final data product, the likelihood the person will be removed increases. The mean false alarm rate also declines as the cameras increase for the corridor LoS sensor. This is also a result of the ANDing of camera. Little change is evident in the mean false alarm rate with a change in the number of cameras when a desk is present, unless there are a large number of people. However, a large increase in the false alarm rate occurs when any camera is added to the terminal without a desk. The

mean precision for the LoS sensor improves as cameras are added. This is also expected as fusion with additional cameras trims the location hypotheses of people.

D. Recommendations for Future Research

1) Further Characterization of LoS Sensor

Further characterization of the LoS sensor performance is needed. Receiver Operating Characteristic curves for each parameter of the sensor would help to optimize design criteria. In addition, the optical design and requirements of the LoS sensor should be studied to optimize performance while determining practical distance, angle, and operating speed limitations of the sensor.

Micro-Electro Mechanical Systems (MEMS) might be used to focus and direct light at both the light source and sensor. MEMS may also be used at the light source to focus the direction the light beam is transmitted, and alter that direction to maximize light transmission towards the light sensor that is synchronized to sense it. MEMS in a light sensor could focus the direction from which the sensor is observes light, and alter that direction such that maximum observation occurs with the light source it is synchronized to sense. If sufficient focus is achieved at the light source and/or sensor, multiple light beams may sweep the light sensors concurrently.

A prototype of the LoS sensor should be built [25]. A small-scale LoS sensor could be used to test its ability to observe toy figures moving within the sensed region. Another possibility is to build a larger prototype in a location with corridors and test its ability to observe actual people and objects. Algorithms used for location determination by the LoS sensor should also be implemented in a Field Programmable Gate Array

(FPGA) [25]. Measurements from either system could improve the simulation model of the LoS sensor.

The sensor network system in which the LoS is integrated should also be evaluated. Image processing algorithms should be exploit the information derived from the LoS sensor data. These algorithms could also be implemented in an FPGA and evaluated with a prototype LoS sensor. Because the LoS sensor may also be used independently of a camera system, data structures and communication methods should be designed for the LoS sensor such that it may be fused with non-camera sensors. Tracking algorithms may be designed, modified or evaluated for use with LoS sensor data.

2) Improve performance of LoS Sensor

One way to optimize the LoS sensor is to determine an optimal placement of light sources and sensor for different applications. For example, the simulation used in this research uses equally spaced lights along a line that are co-planar to and on an opposing surface of an equal number of equally spaced sensors, also along a line. Other distributions of lights and sensors may exist that improve the performance of the LoS sensor at locating objects without significant loss of coverage area. In addition, computational complexity and/or operating speed may be improved by using a different number of light sensors than light sources. Light sensors and sources may also be placed at other locations within the region to be monitored, such as on building support columns or trash receptacles, which may improve the performance of the sensor and increase sensor coverage.

The computer network performance of the LoS sensor should also be evaluated. For example, performance analysis of a wireless sensor network with multiple cameras

should be compared to a similar system design that uses the LoS sensor. A method for the LoS sensor components, such as a bank of lights or sensors, to automatically discover, test, and synchronize can also be examined. A scheduling algorithm may also be developed that seeks to optimize the efficiency of multiple LoS sensors operating concurrently in the same region. In addition, methods of automatic alignment of the LoS sensors should also be explored.

Another area of improvement for the LoS sensor is performance of the simulation model. Many "brute force" techniques are used in the simulation to ensure correctness of the results at the expense of efficiency. Efficient techniques can now be included, the results of which can be verified by comparison to the current, less efficient model [30]. Areas of the simulation model that can be improved include an optimal method of populating the look-up tables, or designing a computationally efficient method of determining the location of objects without the look-up tables.

VI. Final Thoughts

At the completion of this research, an edition of the IEEE Signal Processing Magazine was published that featured surveillance networks [10]. This shows that this is a very active area of research. Given the recent and continuing growth of location aware systems and the increasing desire for more cameras in video networks, there is a need for an efficient method of determining location information. For many conditions, incorporating a network of Line-of-Sight (LoS) sensors is a means of obtaining that efficiency.

The LoS sensor can be used for location determination in a video-free network, but the video network benefits greatly by making the image processing tasks of multi-camera identification and tracking more efficient.

For areas in which there is little pedestrian traffic and few cameras, the LoS sensor may not be appropriate, as the false alarm rate of the sensor may exceed that of a video only system. However, as the number of people grows large, the false alarm rate of the camera system increases significantly. In such cases, the LoS sensor is robust against further increases in the number of people than is a video only system.

The LoS sensor is best applied to areas in which the light sources and light sensors may be placed opposite each other, such as a corridor. As the overall performance of the LoS sensor declines in the presence of large objects that may block many contiguous light rays, results from the sensor may need to be either augmented or ignored in those specific area.

Bibliography

1. Al-Muhtadi, J., A. Ranganathan, R. Campbell and M. D. Mickunas. "Cerberus: a context-aware security scheme for smart spaces," *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications*, p. 489-496, 2003.

2. Azarbayejani, A. and A. Pentland. "Real-time self-calibrating stereo person tracking using 3-D shape estimation from blob features," *Proceedings of the 13th International Conference on Pattern Recognition*, vol. 3, p. 627-632, 1996.

3. Brooks, R.R., P. Ramanathan, and A. M. Sayeed, "Distributed target classification and tracking in sensor networks," *Proceedings of the IEEE*, vol. 91, iss. 8, p. 1163-1171, 2003.

4. Brooks, R. R., S. S. Iyengar, Gunasekaran S. Seetharaman, R. Kannan, Jamie R. Morrison. "Next Generation Distributed Sensor Networks," *6th Annual ONR Workshop on Collaborative Decision-Support Systems*, Quantico Marine Corps Base VA: Office of Naval Research, (September 2004).

5. Brown, L.G. "A survey of image registration techniques," *ACM Compute. Survey*, vol. 24, iss. 4, p. 325-376, 1992.

6. Challa, S., R.J. Evans, and D. Musicki. "Target tracking - a Bayesian perspective," *14th International Conference on Digital Signal Processing.* vol. 1, p. 437-440, 2002.

7. Chen, W.-P., J.C. Hou, and L. Sha. "Dynamic clustering for acoustic target tracking in wireless sensor networks," *11th IEEE International Conference on Network Protocols*, p. 284-294, 2003.

8. Chong, C.-Y. and S.P. Kumar, "Sensor networks: evolution, opportunities, and challenges," *Proceedings of the IEEE*, vol. 91, iss. 8, p. 1247-1256, 2003.

9. Collins, R.T., Alan J. Lipton, Takeo Kanade, Hironobu Fujiyoshi, David Duggins, Yanghai Tsin, David Tolliver, Nobuyoshi Enomoto, Osamu Hasegawa, Peter Burt and Lamber Wixson. *A System for Video Surveillance and Monitoring*, Pittsburg PA: Carnegie Mellon University, 2000. (CMU-RI-TR-00-12).

10. Liu, K. J. Ray. *IEEE Signal Processing Magazine*, vol. 22, num. 2, (March 2005).

11. Gerbino, P. and M. Ali. "A novel local likelihood approach to data fusion in passive target tracking," *IEE Colloquium on Target Tracking: Algorithms and Applications*, p. 2/1 – 2/4, 1999.

12. Goodridge, Steven George. "Multimedia Sensor Fusion for Intelligent Camera Control and Human-Computer Interaction," Raleigh NC: North Carolina State University, 1997.

13. Gordon, N. and D. J. Salmond. "Aspects of target tracking: problems and techniques," *IEE Colloquium on Target Tracking and Data Fusion*, p. 282, 1998.

14. Hazas, M., J. Scott, and J. Krumm. "Location-aware computing comes of age," *Computer*, vol. 37, iss. 2, p. 95-97, 2004.

15. Hills, Rob. "Sensing for Danger," *Science and Technology Review*, p. 11-17, (July/August 2001).

16. Jeong, H. and Jeong-Ho Park, "Multiple target tracking using constrained MAP data association," *Electronics Letters*, vol. 35, iss. 1, p. 25-26, 1999.

17. Katsman, I., A. Bruckstein, R. J. Holt, E. Rivlin. "Judging distance by motion-based visually mediated odometry," *IEEE/RSJ International Conference Proceedings on Intelligent Robots and Systems*, vol. 2, p. 1357-1362, 2003.

18. Katsman, I. and E. Rivlin. "The mantis head camera (why the praying mantis is so good at catching its prey)," *12th International Conference Proceedings on Image Analysis and Processing*, p. 612-617, 2003.

19. Kejun, Zhang, Su Jianbo. "General Software Architecture for Multi-sensor Information Fusion System," *Proceedings of the 5th World Congress on Intelligent Control and Automation*, p. 4640-4644, (June 2004).

20. Leonhardi, A. and K. Rothermel. "Architecture of a large-scale location service," *Proceedings of 22nd International Conference on Distributed Computing Systems*, p. 465-466, 2002.

21. Lou, R. C. and M. G. Kay, "Data Fusion and Sensor Integration: State-of-the-Art 1990's," *Data Fusion in Robotics and Machine Intelligence*, p. 7-135, 1992.

22. Mort, N. and P. Prajitno. "A multisensor data fusion-based target tracking system," *IEEE International Conference on Industrial Technology*, vol. 1, p. 427-432, 2002.

23. Patterson, C.A., R.R. Muntz, and C.M. Pancake. "Challenges in location-aware computing," *IEEE Pervasive Computing*, vol. 2, iss. 2, p. 80-89, 2003.

24. Salmond, D.J. "Mixture reduction algorithms for target tracking," *IEE Colloquium on State Estimation in Aerospace and Tracking Applications*, p. 7/1 – 7/4, 1989.

25. Salmond, D.J. "Target tracking: introduction and Kalman tracking filters," *IEE Workshop on Target Tracking: Algorithms and Applications*, vol. 2, p. 1/1-1/6, 2001.

26. Seetharaman, Gunasekaran S. *Personal Communication*, Dayton OH: Air Force Institute of Technology, (February 2005).

27. Seetharaman, Gunasekaran S., Ha V. Le, S. S. Iyengar and H. Legananthraj. "SmartSAM: A Multisensor Network Based Framework for Video Surveillance and Monitoring," Lafayette LA: University of Louisiana Center for Advanced Computer Studies, 2004.

28. Seetharaman, Gunasekaran S., Ha V. Le, S. S. Iyengar, and H. Legananthraj. "A Motion compensated super resolution imaging technique," *Third International Workshop on Digital and Computational Video*, p. 35-42, 2002.

29. Want, R. and A. Hopper, "Active badges and personal interactive computing objects," *IEEE Transactions on Consumer Electronics*, vol. 38, iss. 1, p. 10-20, 1992.

30. Wood, Christopher. *Personal Communication*, Dayton OH: Air Force Institute of Technology, (February 2005).

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From – To)*
21-03-2005	Master's Thesis	March 2004 – March 2005

4. TITLE AND SUBTITLE

A Line-Of-Sight Sensor Network for Wide Area Video Surveillance: Simulation and Evaluation

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Morrison, Jamie R., First Lieutenant, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(S)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
Building 641
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/GCE/ENG/05-05

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

AFRL/HECV
Attn: Dr. Paul Havig (email: Paul.Havig@wpafb.mil)
2255 H. Street
WPAFB OH 45433-7022

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Substantial performance improvement of a wide area video surveillance network can be obtained with addition of a Line-of-Sight sensor. The research described in this thesis shows that while the Line-of-Sight sensor cannot monitor areas with the ubiquity of video cameras alone, the combined network produces substantially fewer false alarms and superior location precision for numerous moving people than video.

Recent progress in fabrication of inexpensive video cameras have triggered a new approach to wide area surveillance of busy areas such as modeling an airport corridor as a distributed sensor network problem. The computation and communication to establish image registration between the cameras grows rapidly as the number cameras increases. Computation is required to detect people in each image; establish a correspondence between people in two or more images; compute exact 3-D position from each corresponding-pair; and temporally track targets in space-and-time. Substantial improvement can be obtained with addition of a Line-of-Sight sensor as a location detection system to decoupling the detection, localization, and identification subtasks. That is, if the 'where' can be answered by a location detection system, the 'what' can be addressed by the video most effectively.

15. SUBJECT TERMS

VIDEO NETWORKS, VISUAL SURVEILLANCE, SENSOR FUSION, OPTICAL TRACKING, AUTOMATIC TRACKING, IMAGE PROCESSING

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
REPORT	**ABSTRACT**	**c. THIS PAGE**			Dr. Guna Seetharaman
U	U	U	UU	168	**19b. TELEPHONE NUMBER** *(Include area code)* (937) 255-3636, ext 4612; e-mail: Guna.Seetharaman@afit.edu